Time Series Analysis of Remotely sensed data & accuracy assessment

Mutlu Ozdogan
ozdogan@wisc.edu

TAT-2017 Pecs, HU
June 19, 2017

Outline

- Definitions
- Ways to use time series analysis
- Examples
- New methods
- Accuracy assessment
- Wrap-up

Take a break with images

We will take a break and play the

"*guess where that is*"

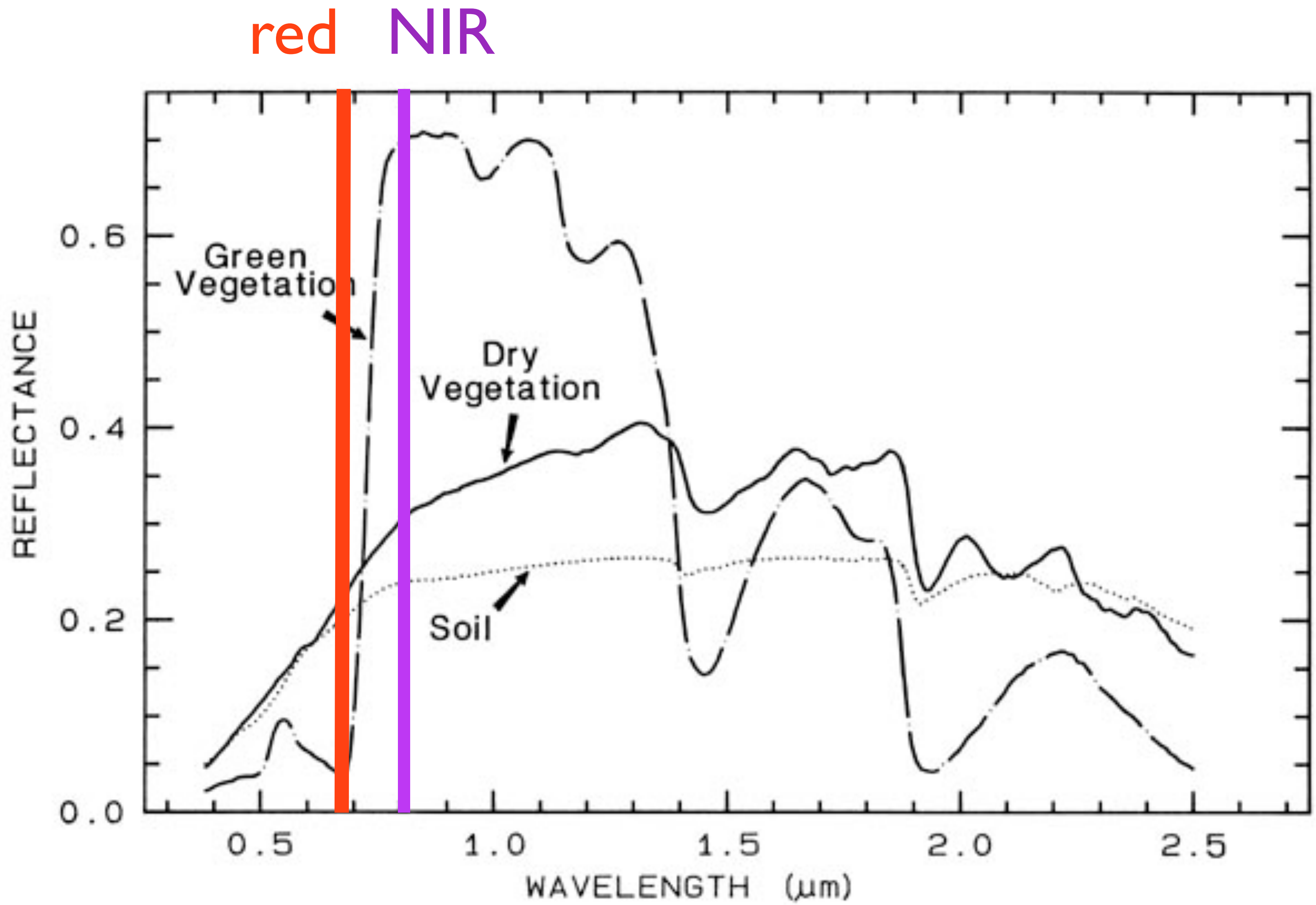
game with images between different parts of the lecture …

What is time series?

- In statistics, signal processing and mathematical finance, a time series is a sequence of data points, measured typically at successive times spaced at uniform or non-regular time intervals
- Examples of time series are the daily closing value of the Dow Jones index or the annual flow volume of the Colorado River at Hoover dam
- In remote sensing, high time frequency observations made regularly at every pixel comprise a time series.
- Remotely sensed data with time series availability is sometimes called “hyper-temporal”

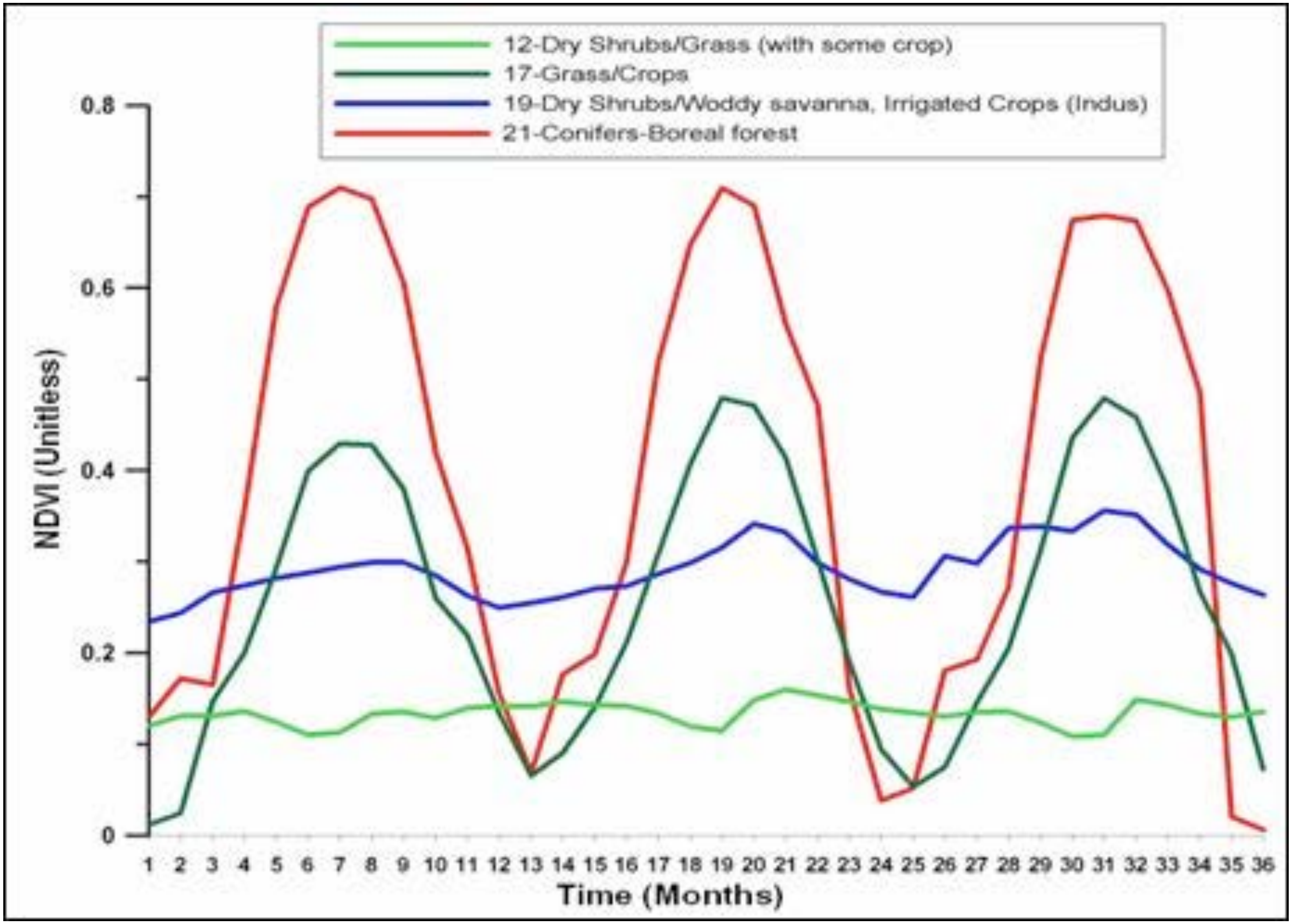
Wikipedia

spectral example

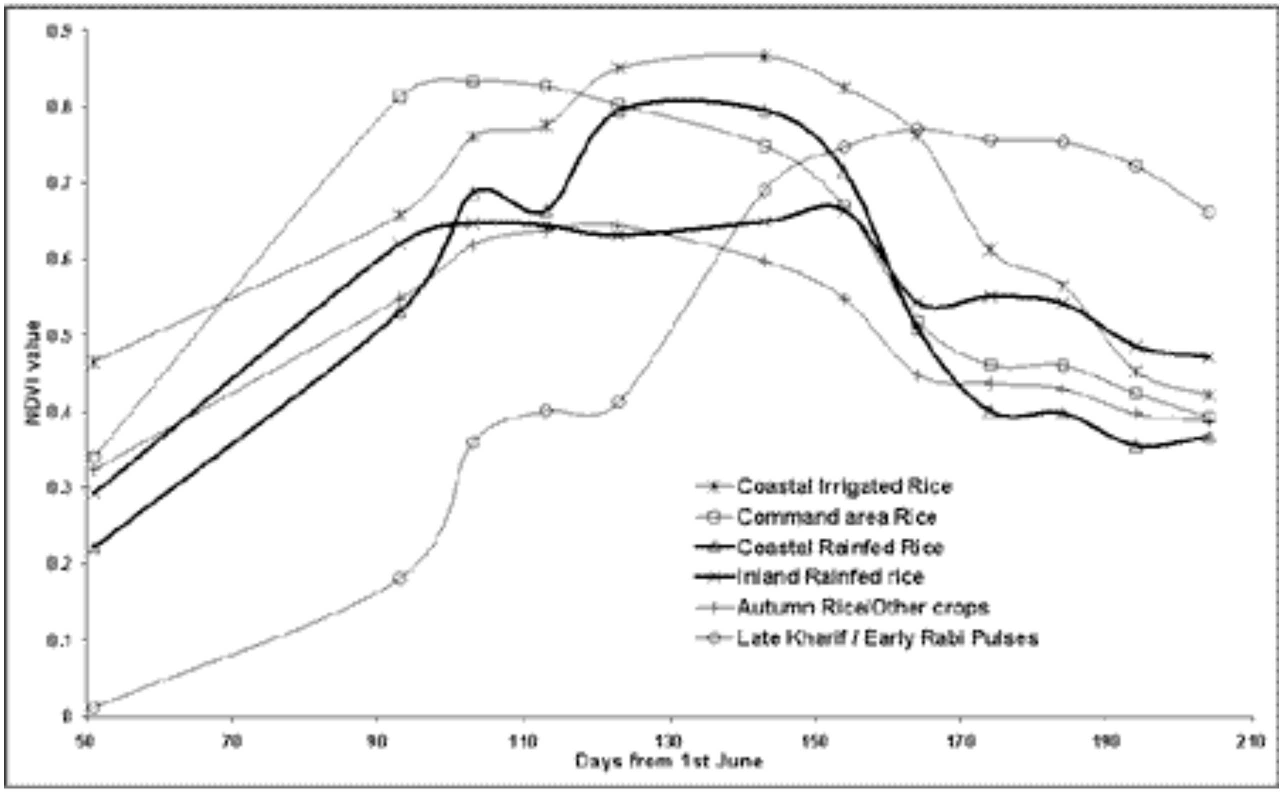


summer image

temporal example



temporal example



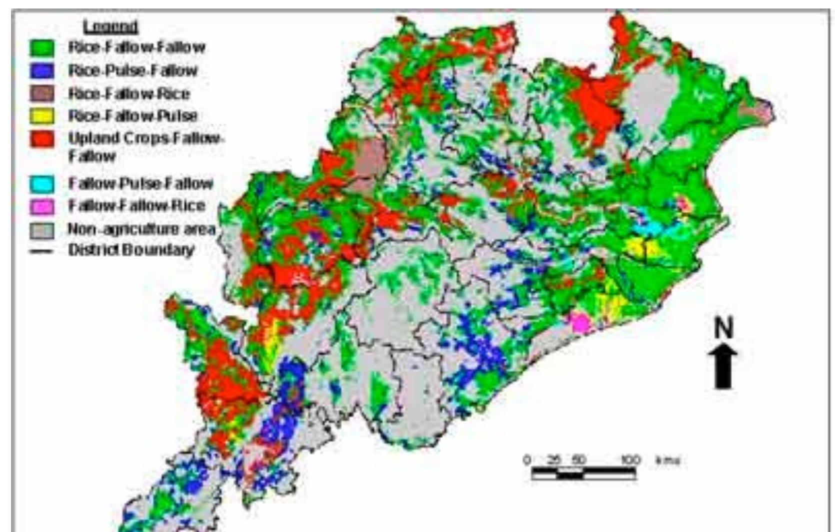


Fig 3: CROP ROTATION MAP OF ORISSA STATE (derived using SPOT VGT data)

Time series analysis

- In remote sensing, time series analysis refers to tools and methods to extract information about the landscape characterized by <u>both</u> spectral <u>and</u> temporal changes

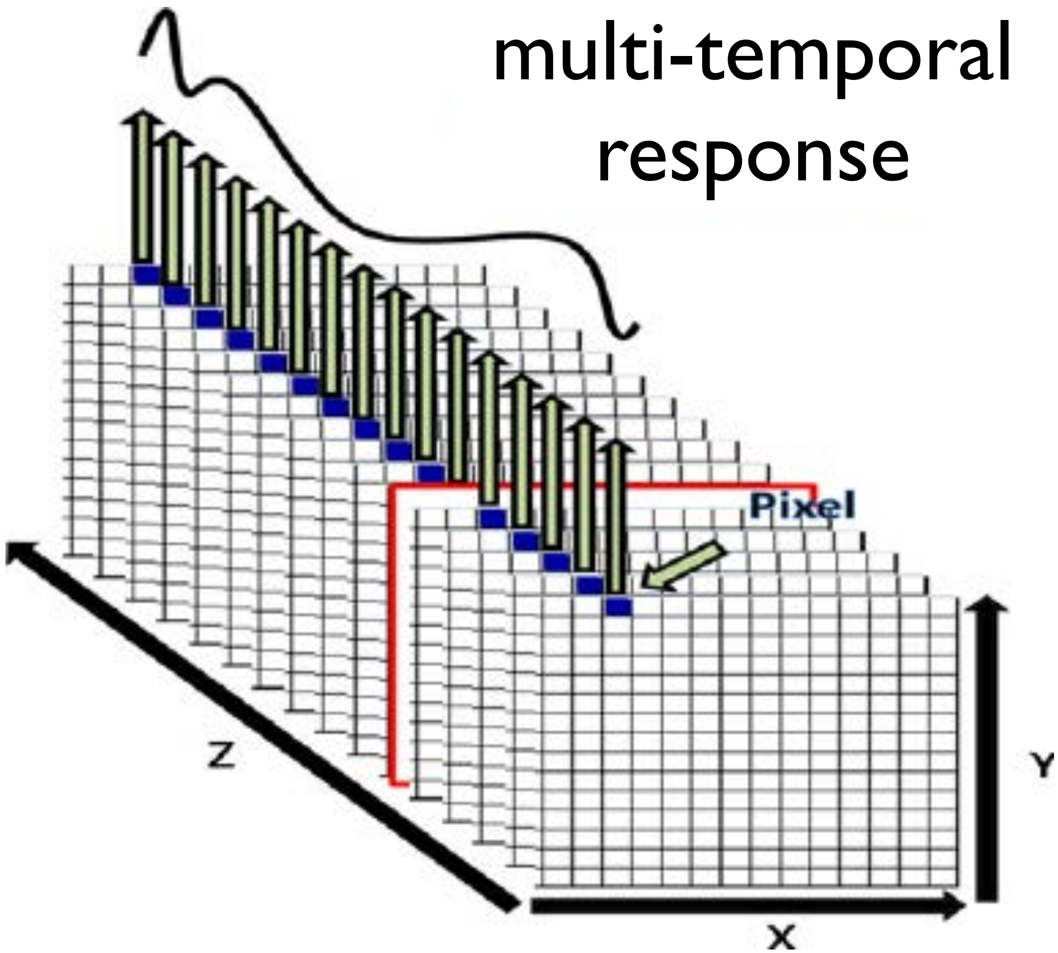
- Unlike most time series analysis, forecasting future events based on known past events is <u>not</u> performed in remote sensing

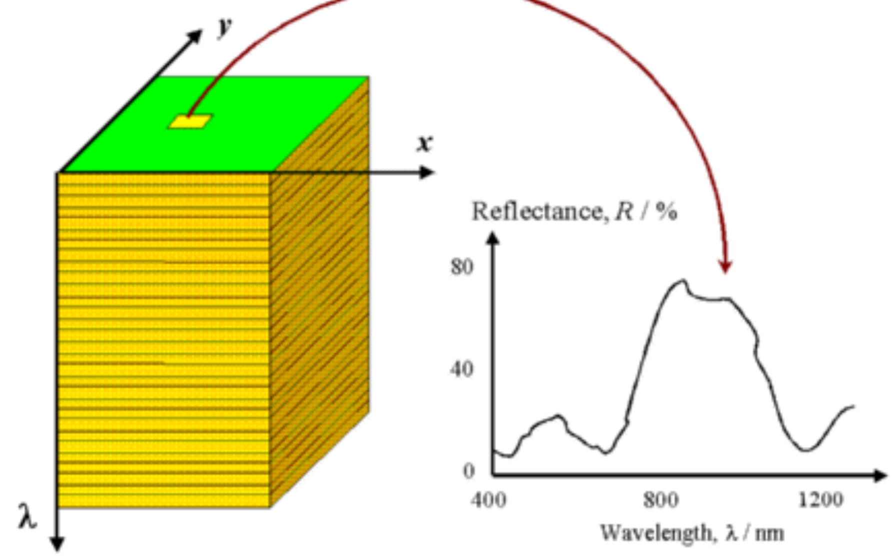
Wikipedia

Time series analysis

- In remote sensing, time series analysis tools are often applied to individual pixels independently (i.e. no interaction between pixels)
- The key is to apply the same method to <u>every</u> pixel which means same application over millions of pixels/locations
- This means many time series methods when applied over a spatial domain (e.g. a country) are computationally expensive - may require a parallel computing environment (think cloud computing)

multi-temporal response

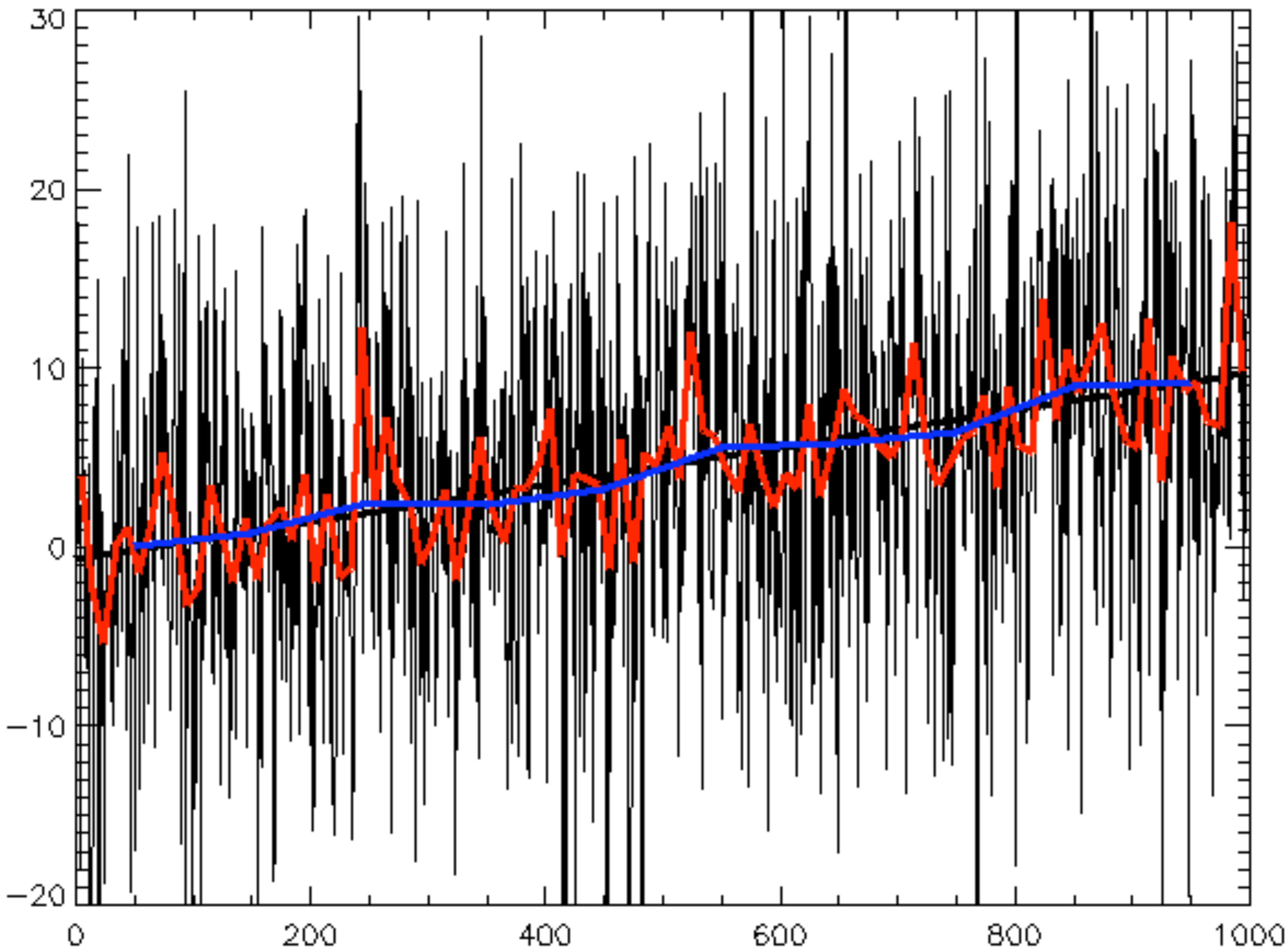




Some terminology

- High Temporal Frequency - Observations made with high temporal repetition - length between two successive observations are small (days to weeks)
- Smoothing - includes methods to "smooth" or remove high-frequency signal such as median smoothing, moving average etc.
- Trend - long-term character of a series of observations

example of a high frequency event (black)
same event smoothed (filtered) (red)
trend of the event over time (blue)



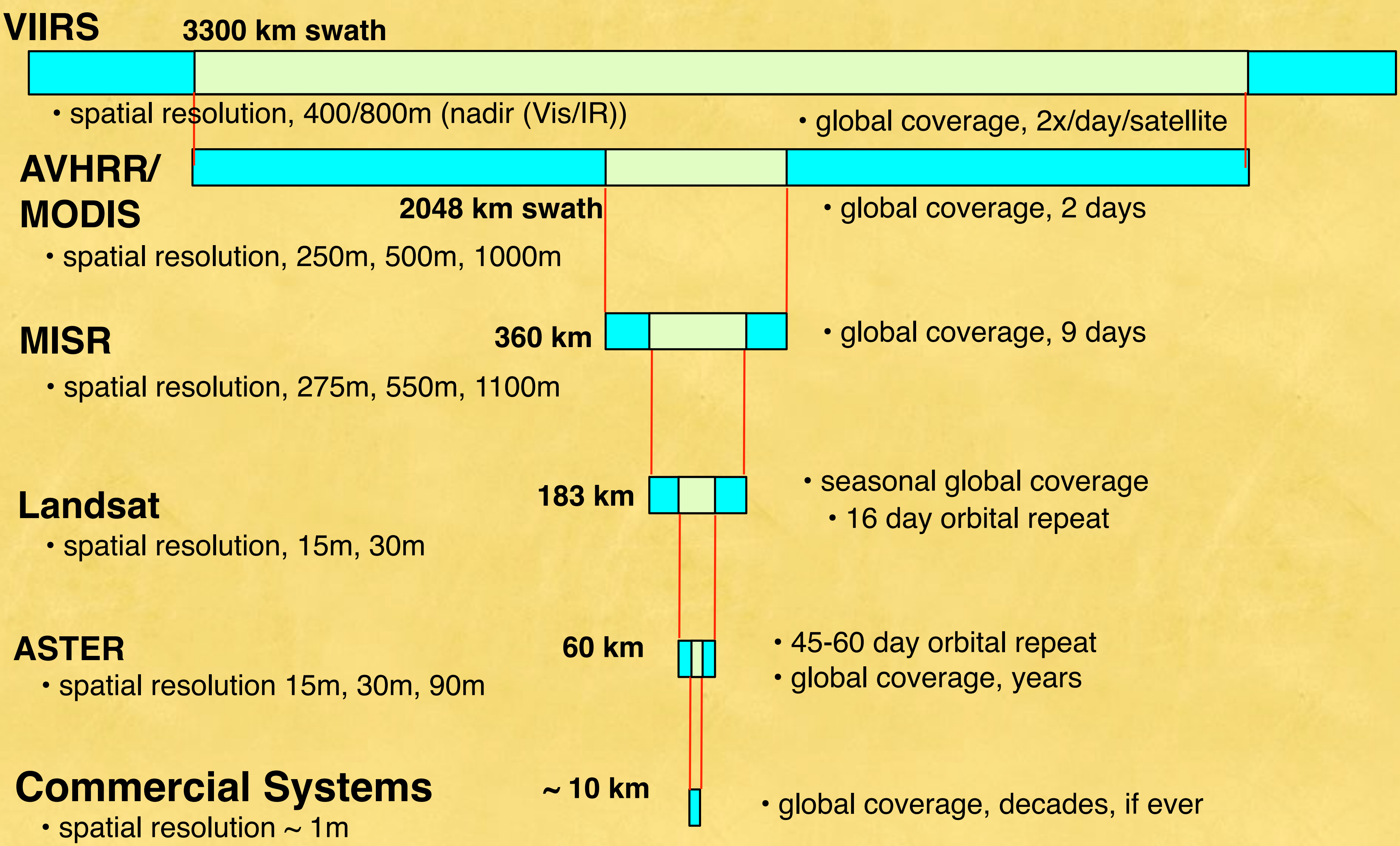
Wikipedia

Dichotomy (Duality) of remote sensing observations in time series analysis

- Data from high time frequency sensors (e.g. MODIS, AVHRR, SPOT Vegetation, Sentinel-3, PROBA-V etc.) are often regularly spaced in time and allow established time series methods to be applied

- Data from (traditionally) time infrequent sensors (e.g. Landsat, Sentinel etc.) are irregularly observed and contain many noise (e.g. clouds) so traditional time series methods may not always apply. Time series data from Landsat-like sensors is often called "dense time stacks"

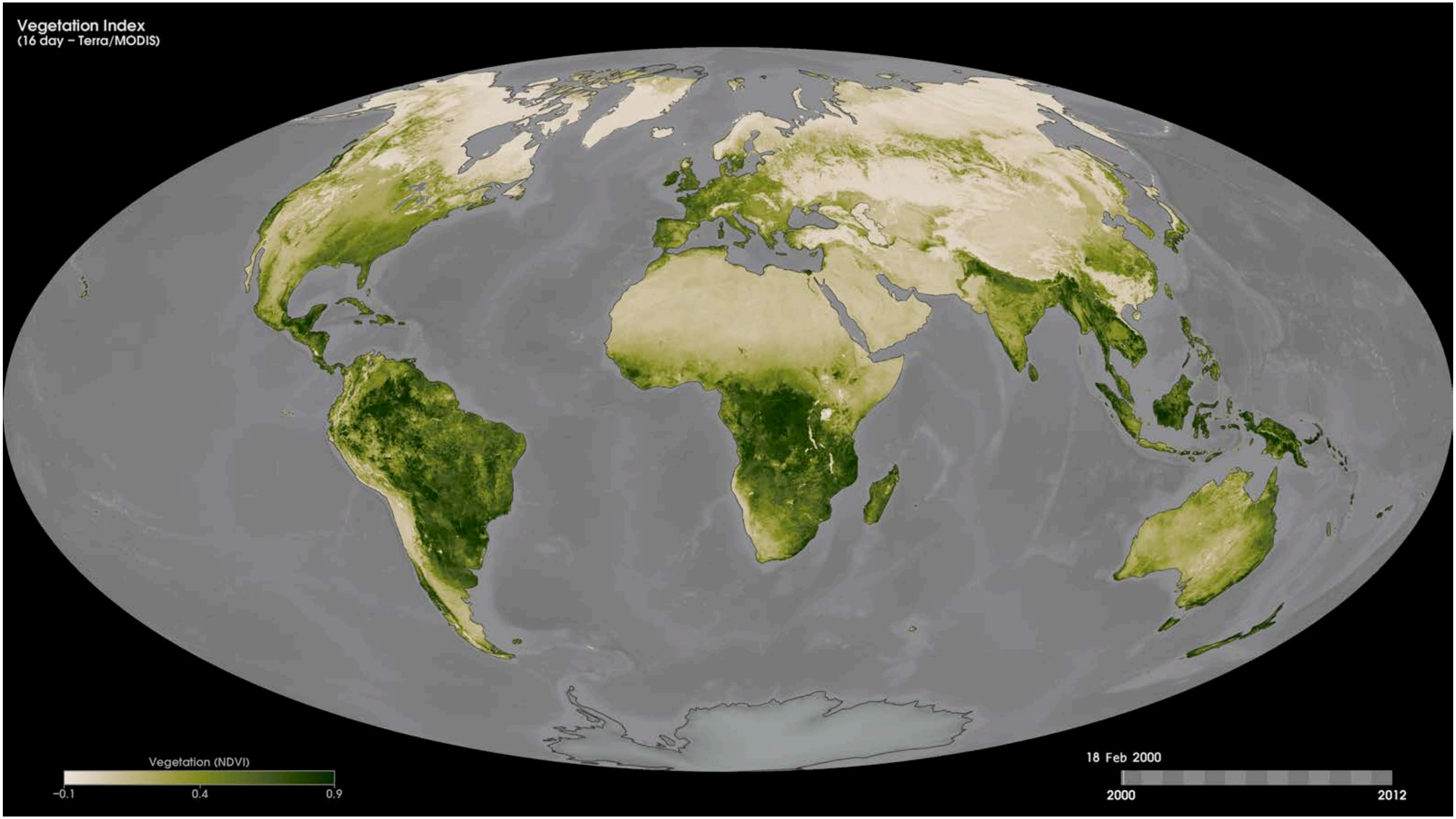
Synergistic Use of Optical Remote Sensing



MODIS vegetation index product

https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod13q1

MODIS NDVI

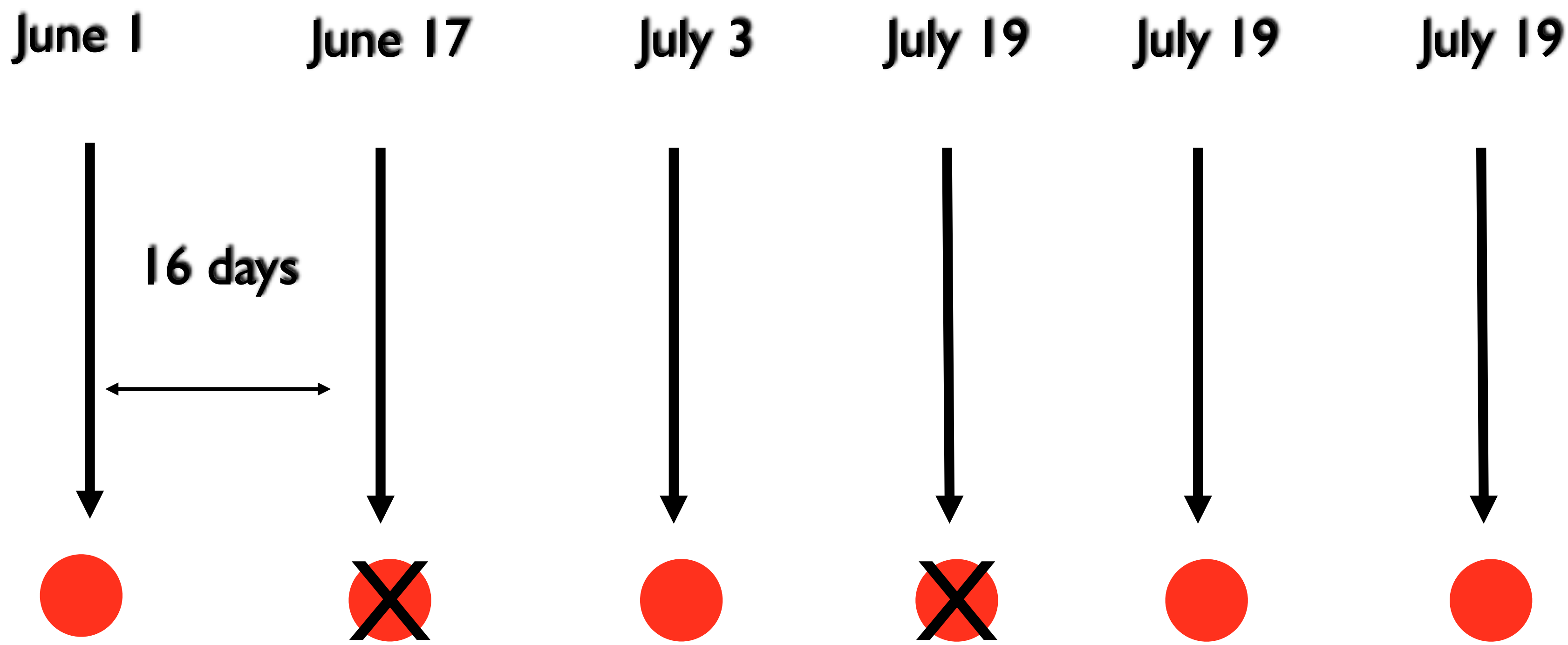


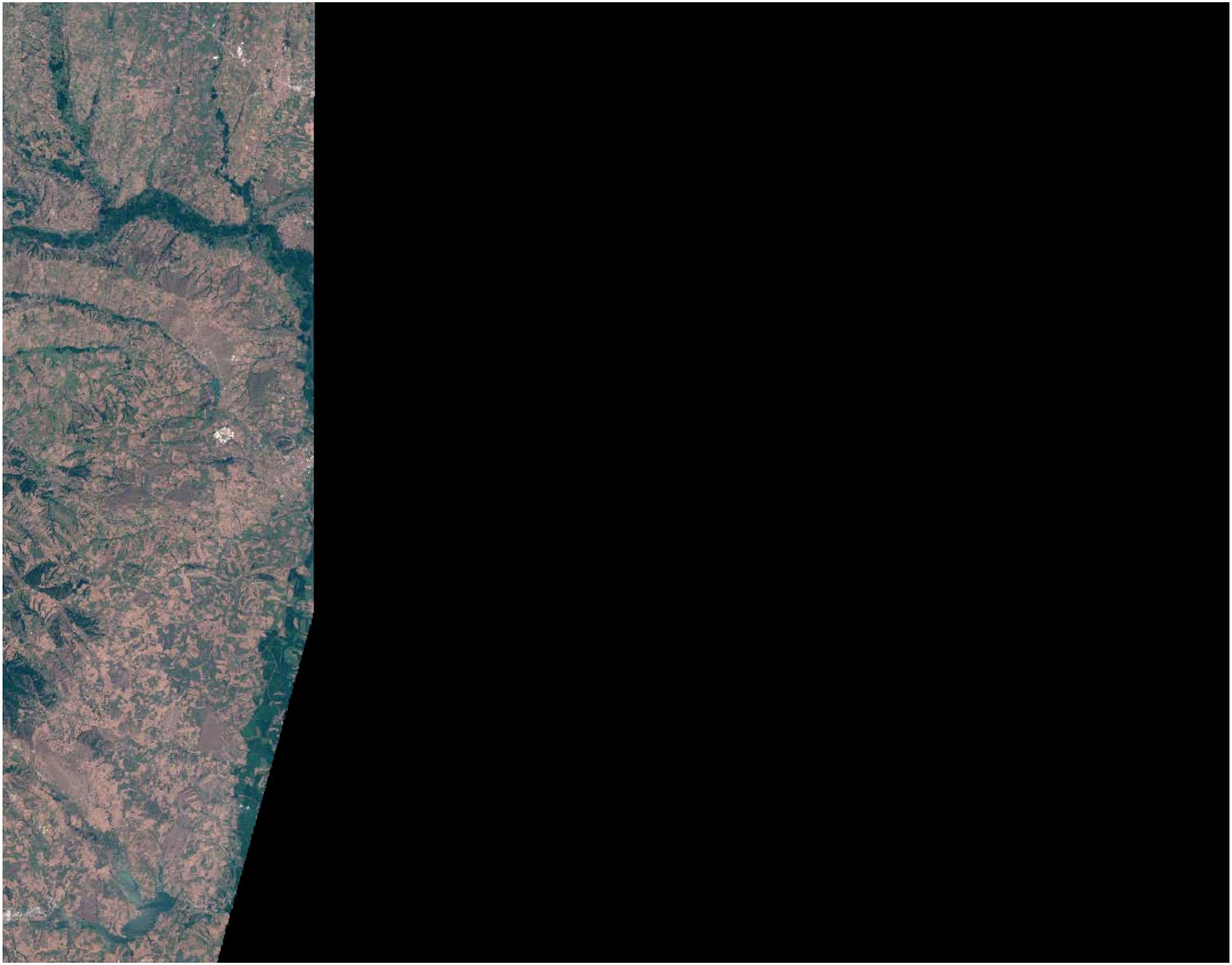
Vegetation Index
(16 day – Terra/MODIS)
Vegetation (NDVI)
−0.1
0.4
0.9
18 Feb 2000
2000
2012

Landsat acquisitions

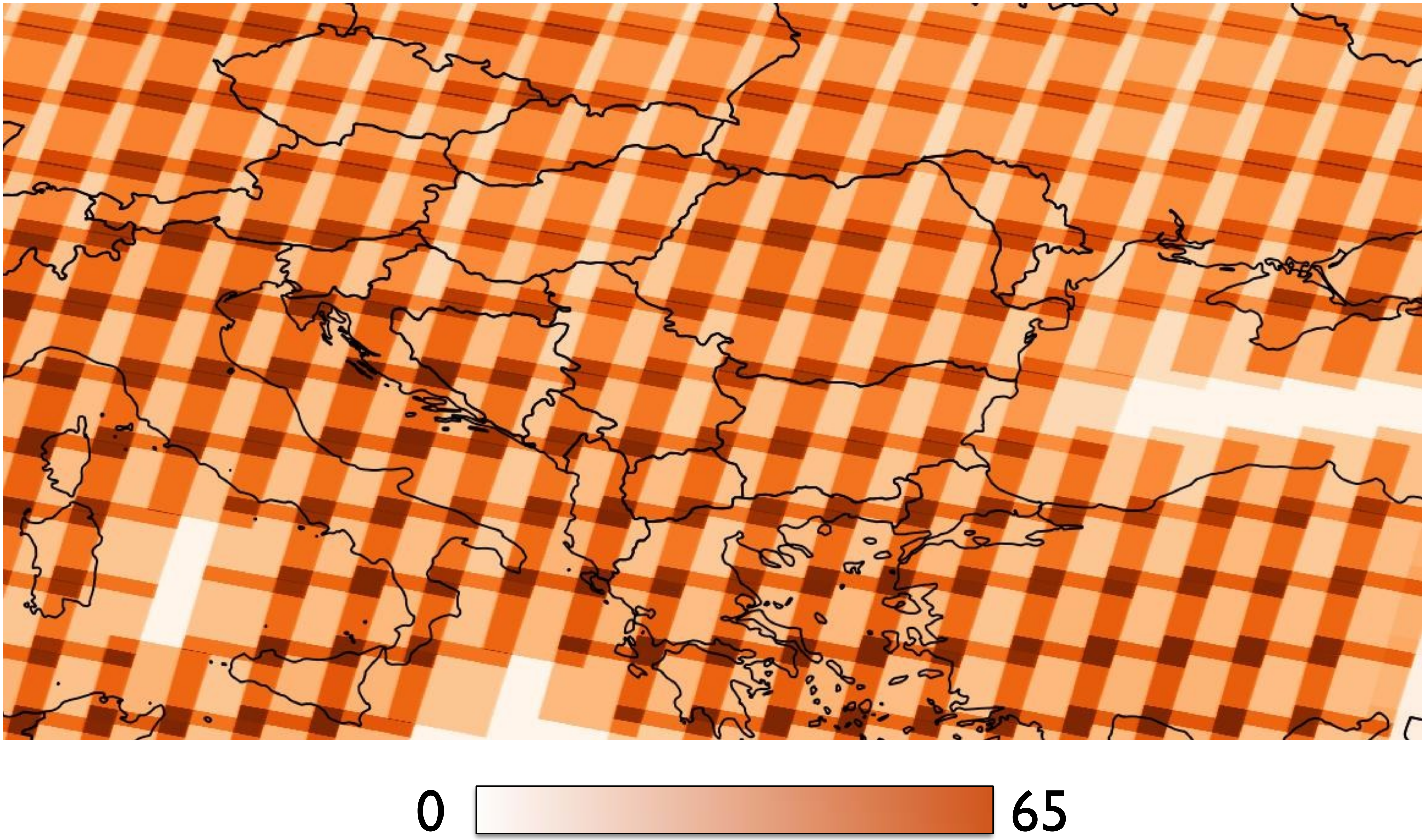
Temporal Resolution

Landsat Data Acquisition

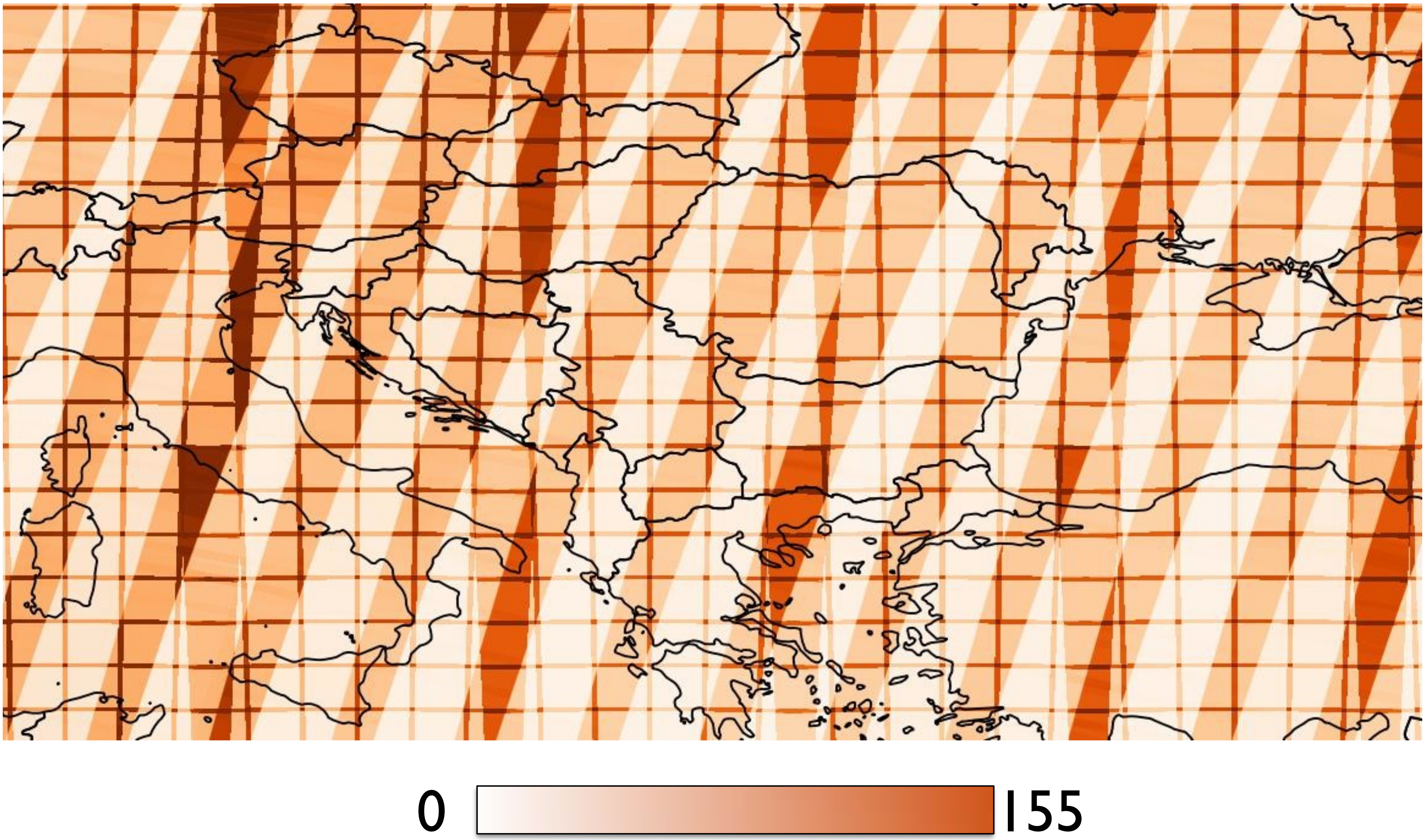


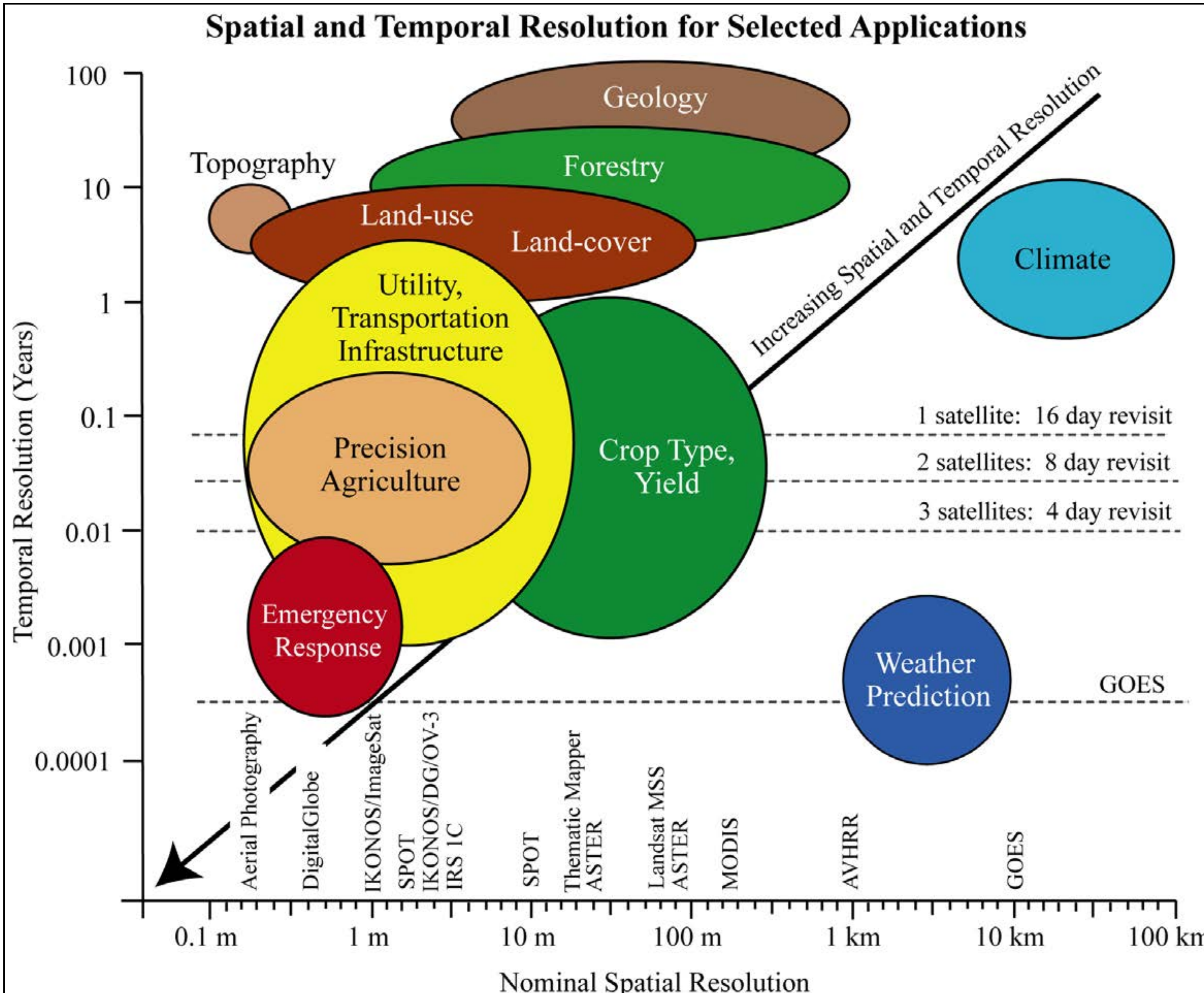


Landsat acquisition density

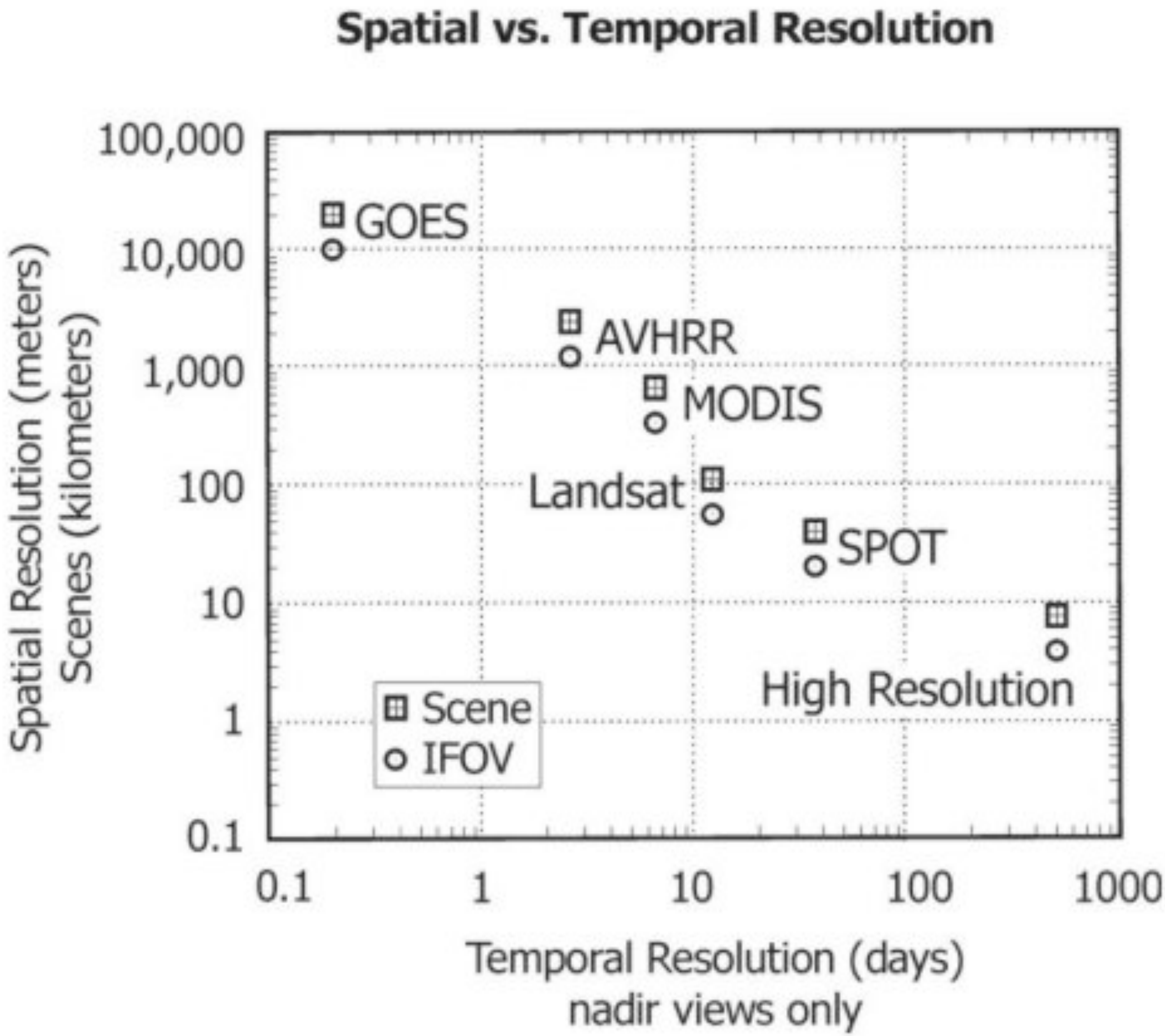


Sentinel-2 acquisition density





There are spatial and temporal resolution considerations that must be made for certain remote sensing applications.


Spatial vs. Temporal Resolution
100,000
10,000
1,000
100
10
1
0.1
Spatial Resolution (meters)
Scenes (kilometers)
GOES
AVHRR
MODIS
Landsat
SPOT
High Resolution
Scene
IFOV
0.1
1
10
100
1000
Temporal Resolution (days)
nadir views only

Time series analysis methods

Pre-processing

- gap fill
- averaging (smoothing)
- smart filtering

Information extraction

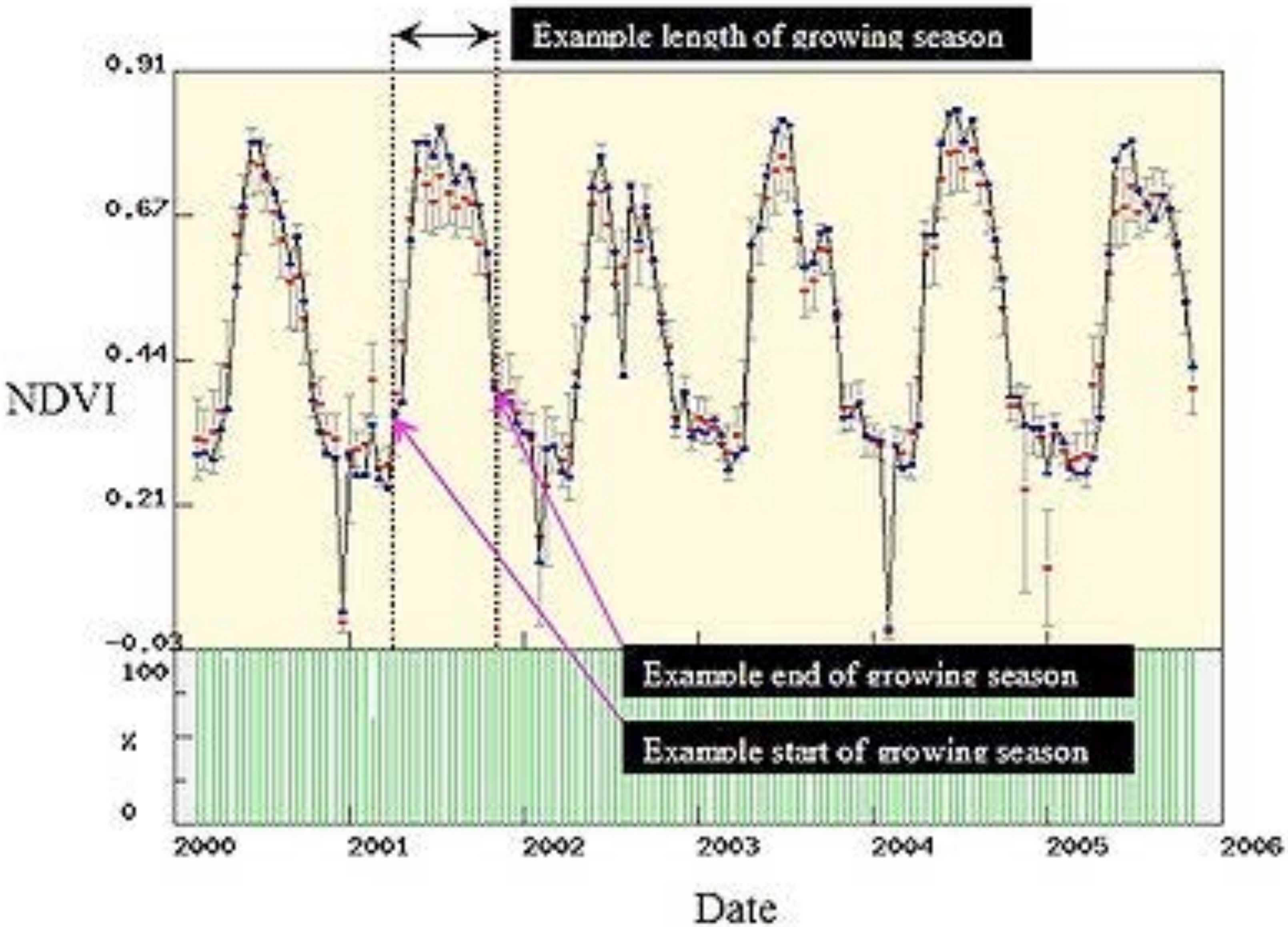
- Classification
- Phenological metrics
- Fourier analysis
- temporal unmixing
- many more....

pre-processing

Pre-processing

- As with many other remotely sensed observations, the time series data tends to be noisy and must be "pre-processed" before any information can be extracted
- The key to successful pre-processing is to find the sweet spot where neither too much information is lost nor too much noise is retained
- In some cases (though rare), the noise you are trying to remove contains the information you need

temporal example



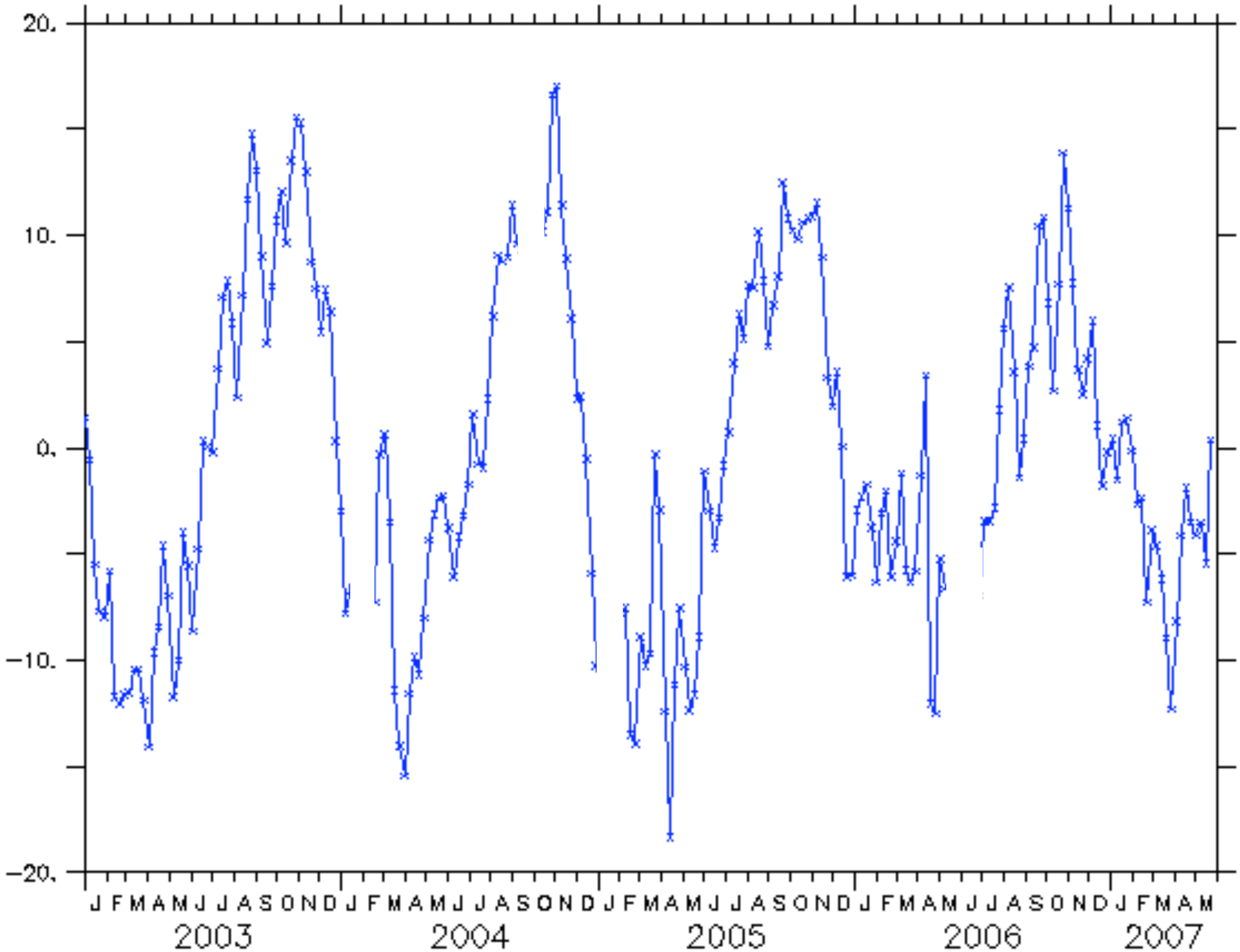
What are the sources of noise?

- atmospheric
  - clouds
  - atmospheric attenuation
  - atmospheric BRDF
- surface related
  - surface BRDF
  - surface cover type
  - phenology
  - differential illumination

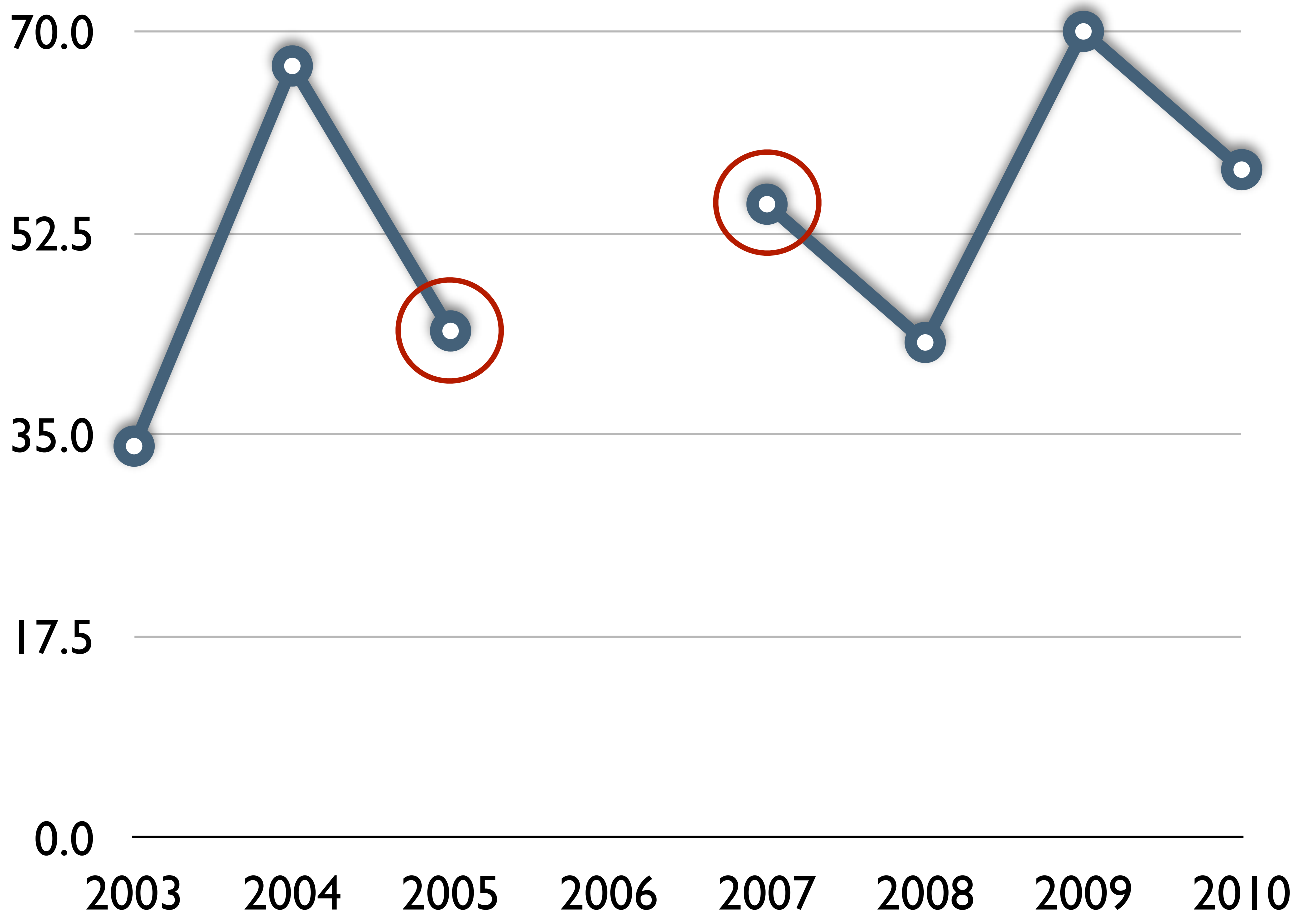
Temporal pre-processing methods

- Gap fill: Simply treat noise as missing data and fill the missing data (or gaps) with either ancillary information or with neighboring information (also called imputation)
- Averaging (smoothing): run a statistical smoothing filter (mean, median, majority etc...) on the time series data - window size determines the magnitude of the effect
- Smart filtering: run a locally adaptive filter that while removing noise, preserves the original shape of the data (e.g. Savitzky-Golay filter)

gap-fill example

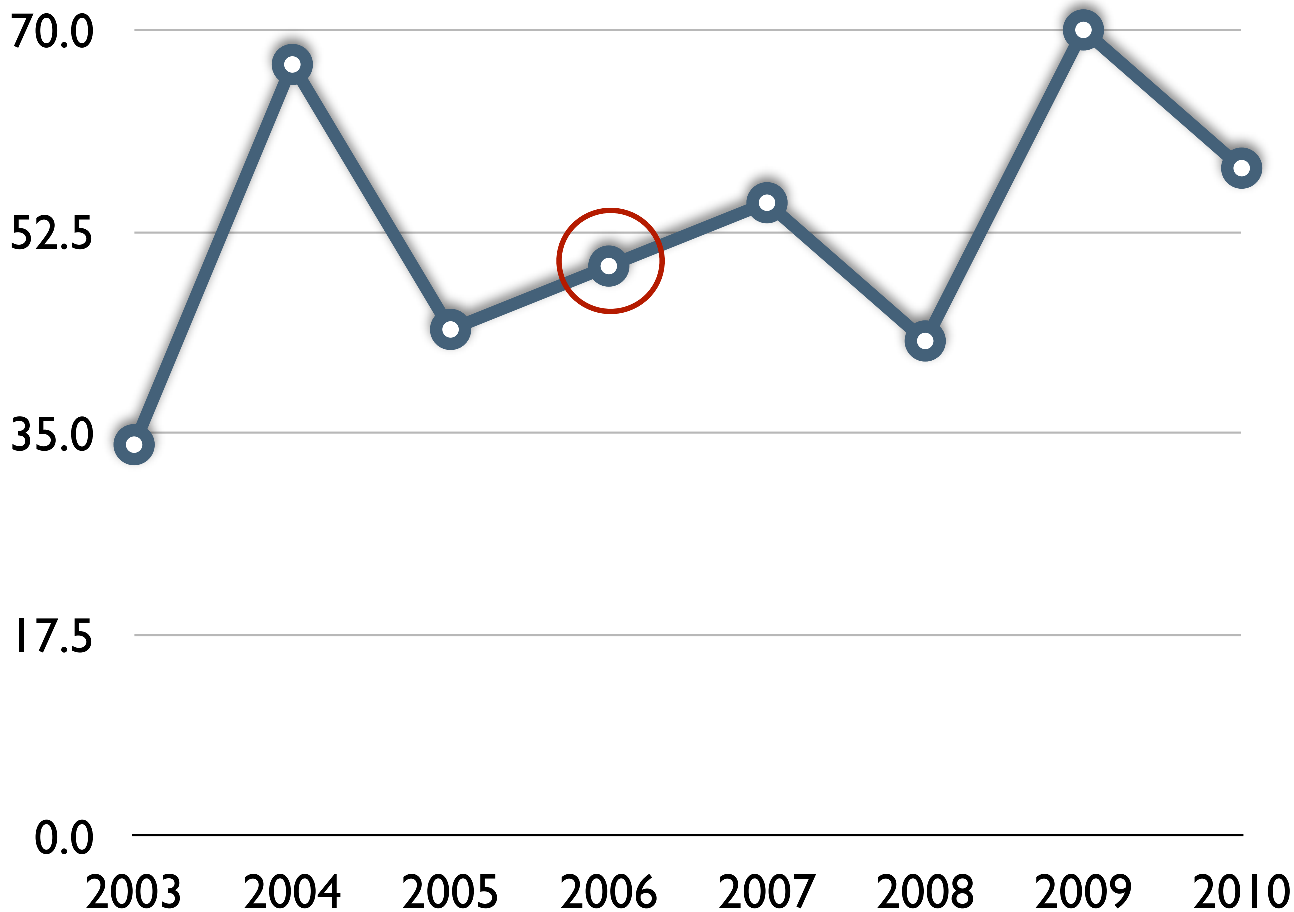


gap-fill example



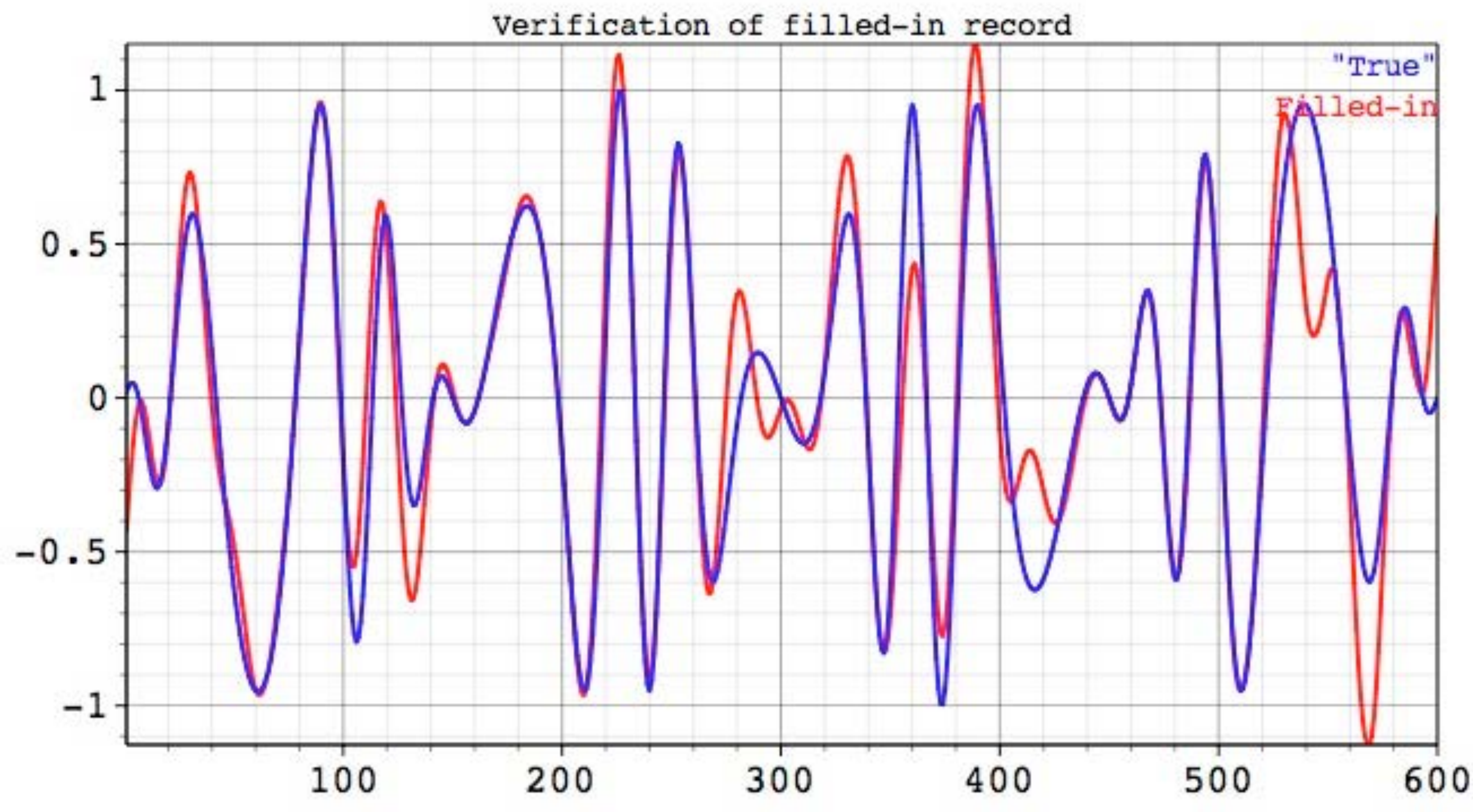
$$Y_t = (Y_{[t-1]} + Y_{[t+1]})/2$$

gap-fill example

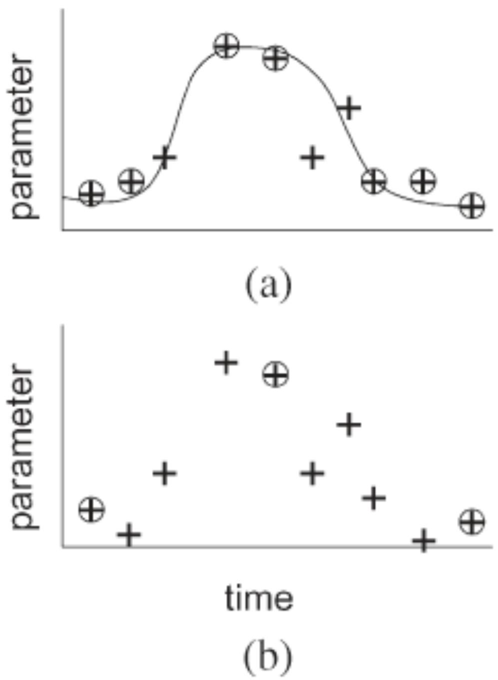


$$Y_t = (Y_{[t-1]} + Y_{[t+1]})/2$$

gap-fill example

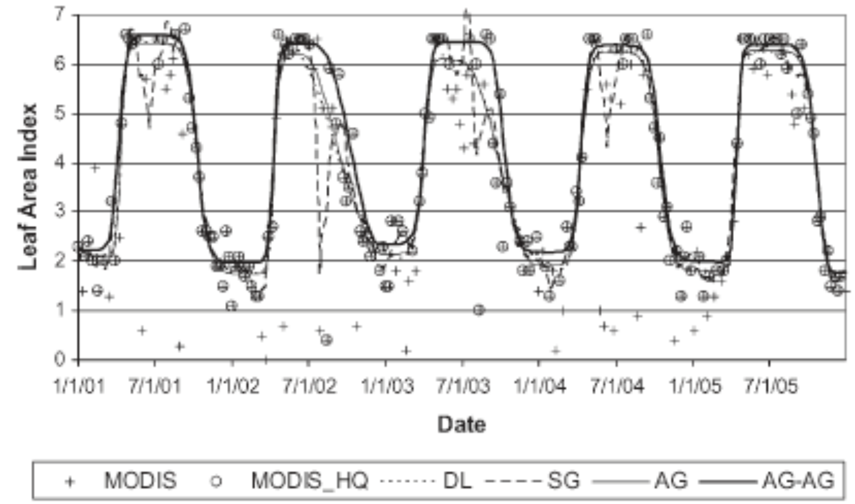


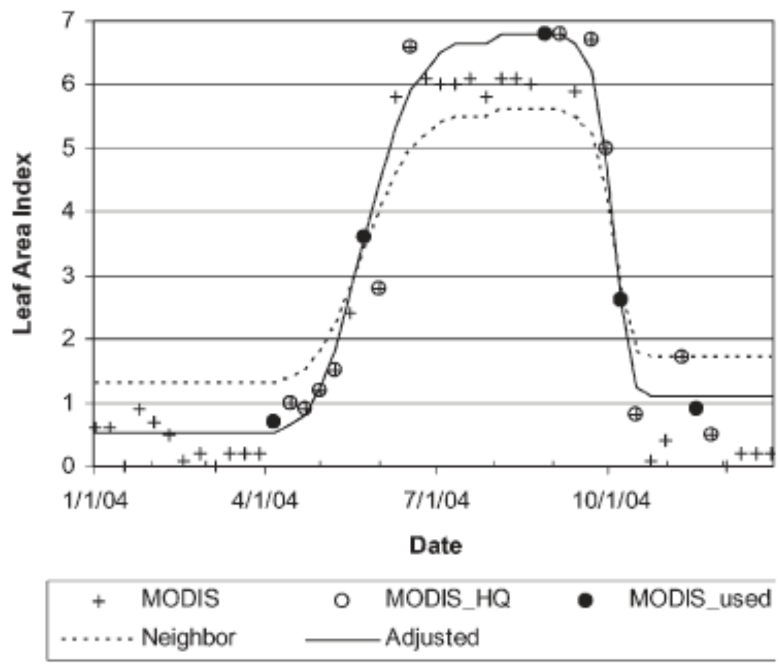
gap-fill example



Case 1
typical situation:
enough high quality
data to adequately fit
curve to the retrievals

Case 2
degenerate situation:
not enough high
quality pixels to
adequately fit curve to
the retrievals

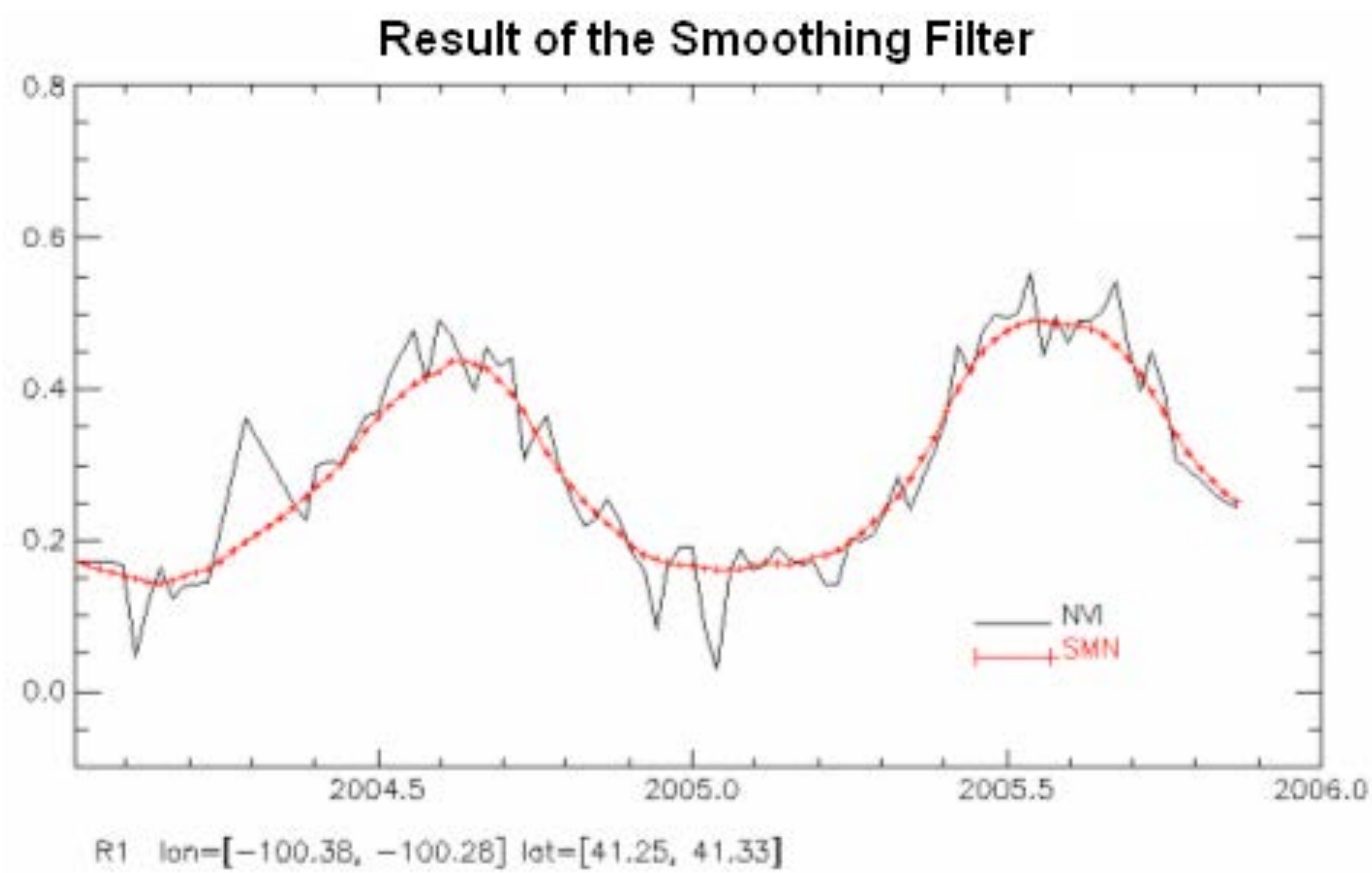




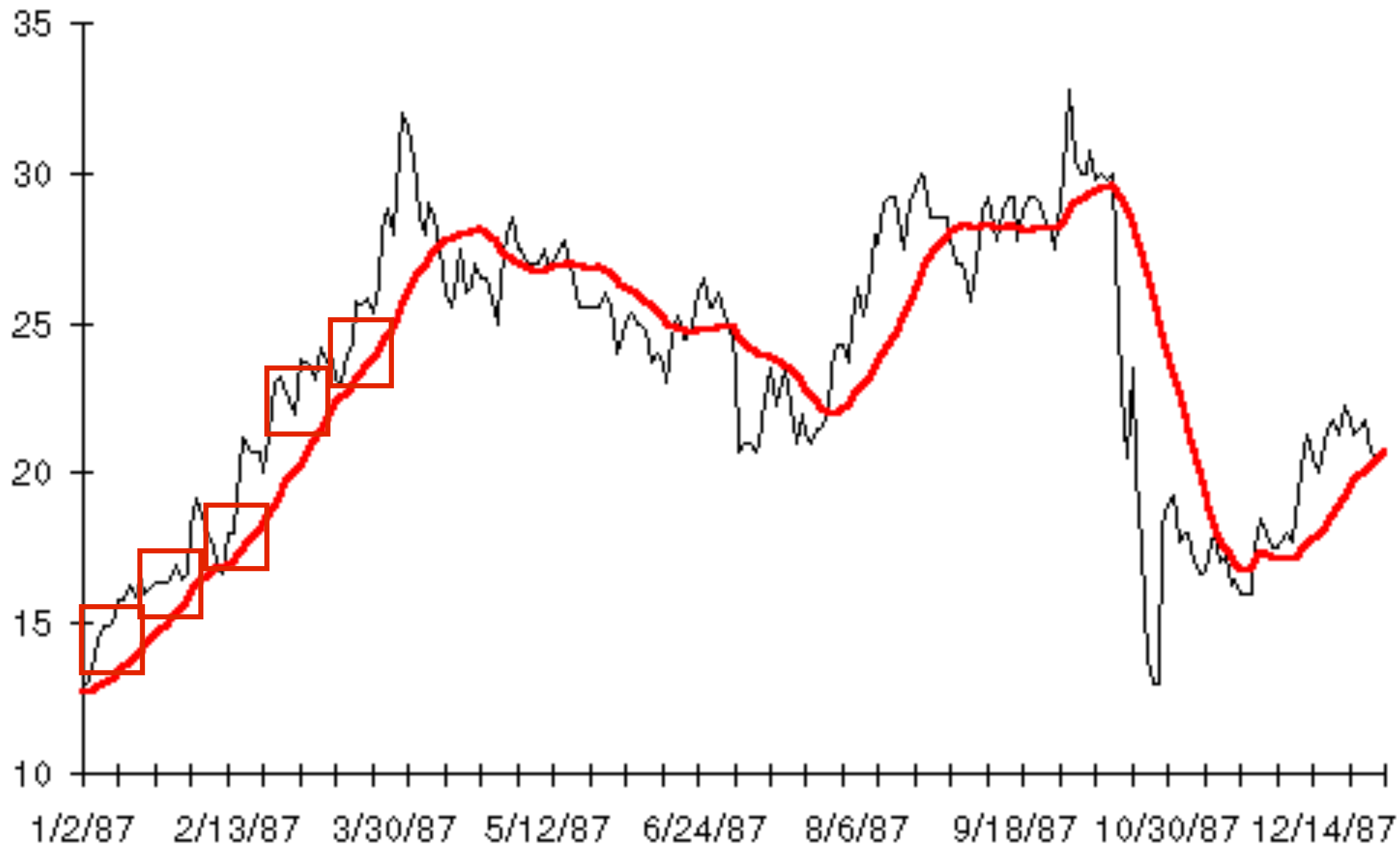
Gao et al (2008)

Smoothing Example

- Mean, median,moving average etc
- By fitting a curve



moving average example

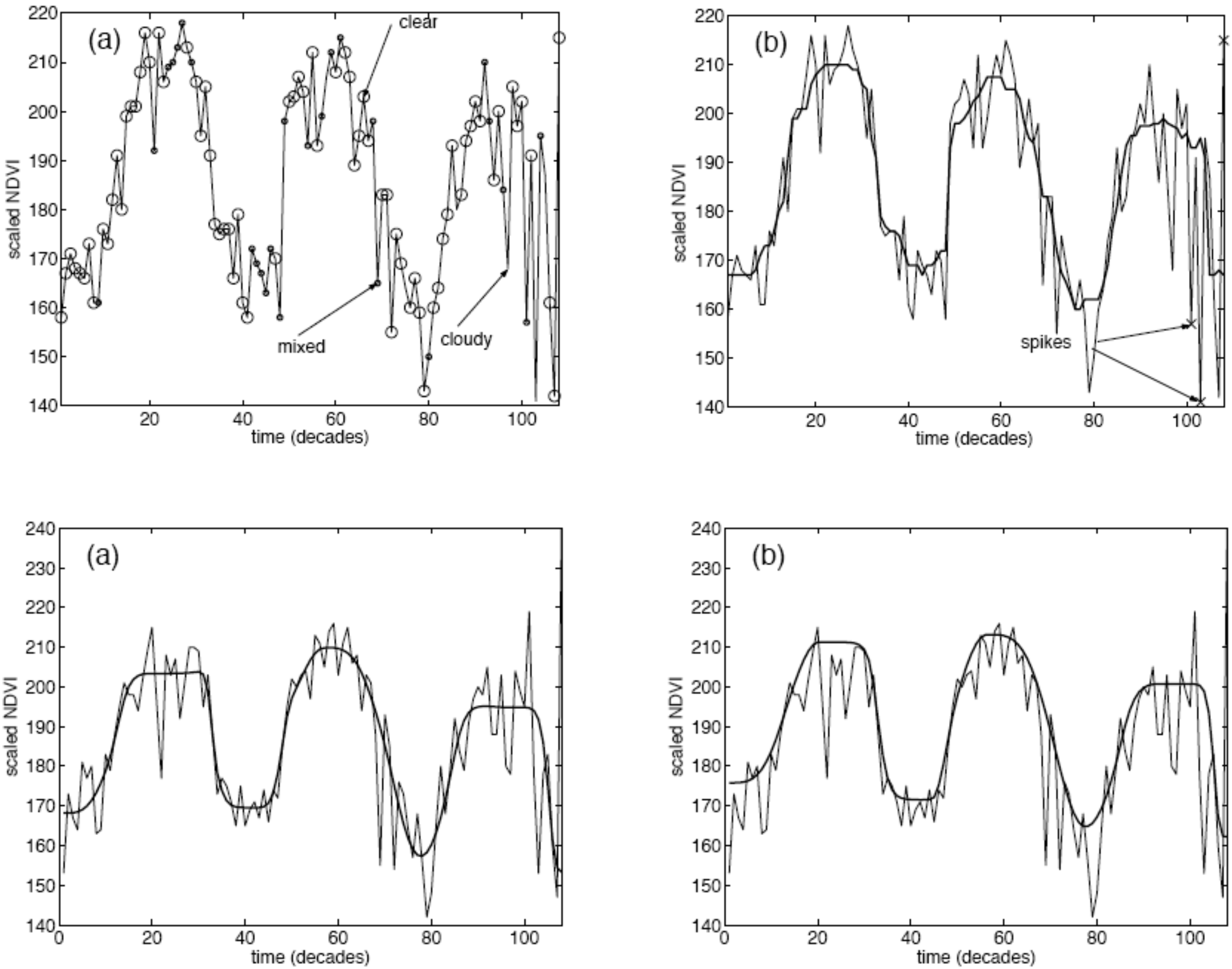


smart filtering example

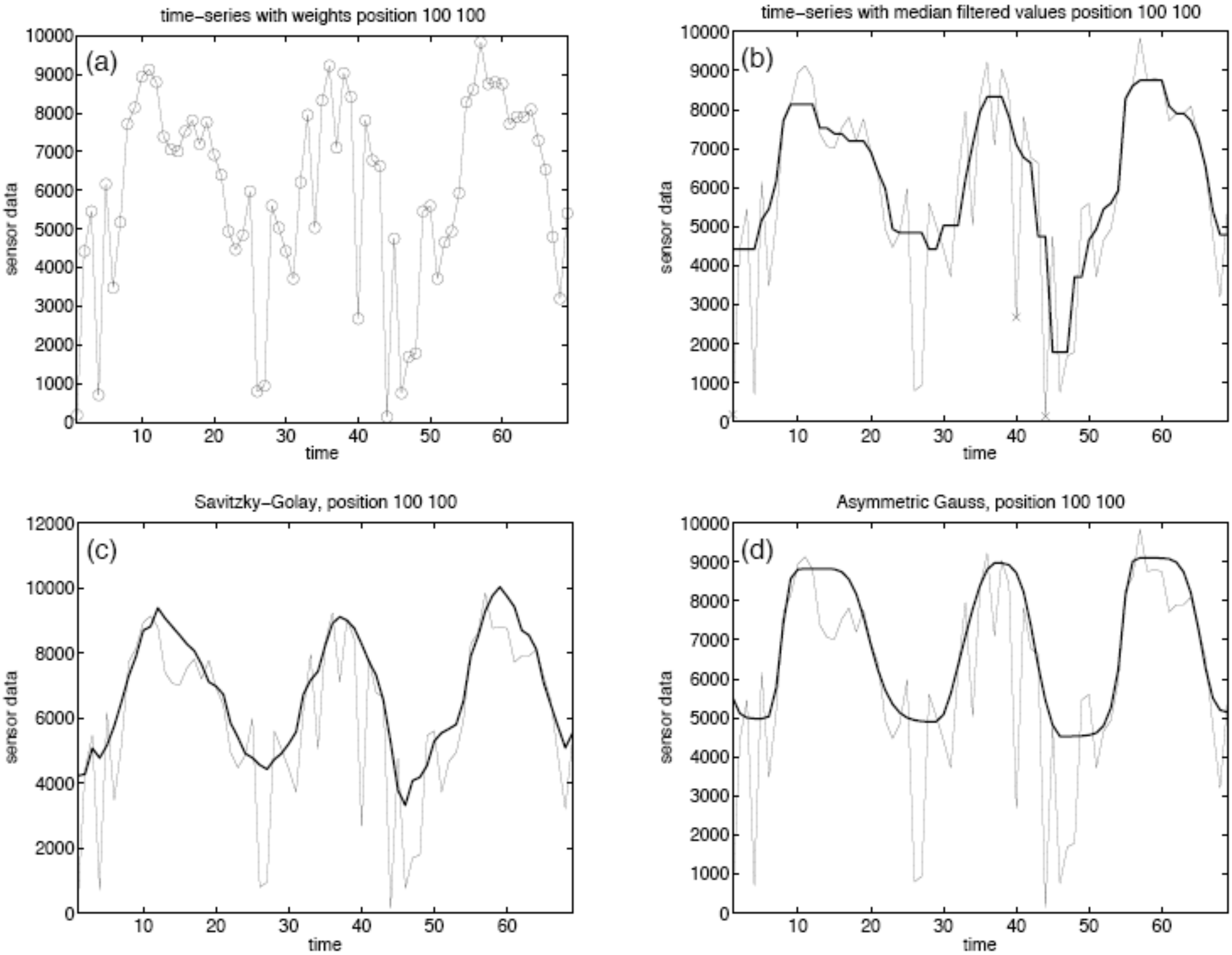
- The Savitzky–Golay smoothing filter is a type of filter first described in 1964 by Abraham Savitzky and Marcel J. E. Golay
- The Savitzky–Golay method essentially performs a local polynomial regression (of degree k) on a series of values to determine the smoothed value for each point
- The main advantage of this approach is that it tends to preserve features of the distribution such as relative maxima, minima and width, which are usually 'flattened' by other adjacent averaging techniques (like moving averages, for example).

Wikipedia

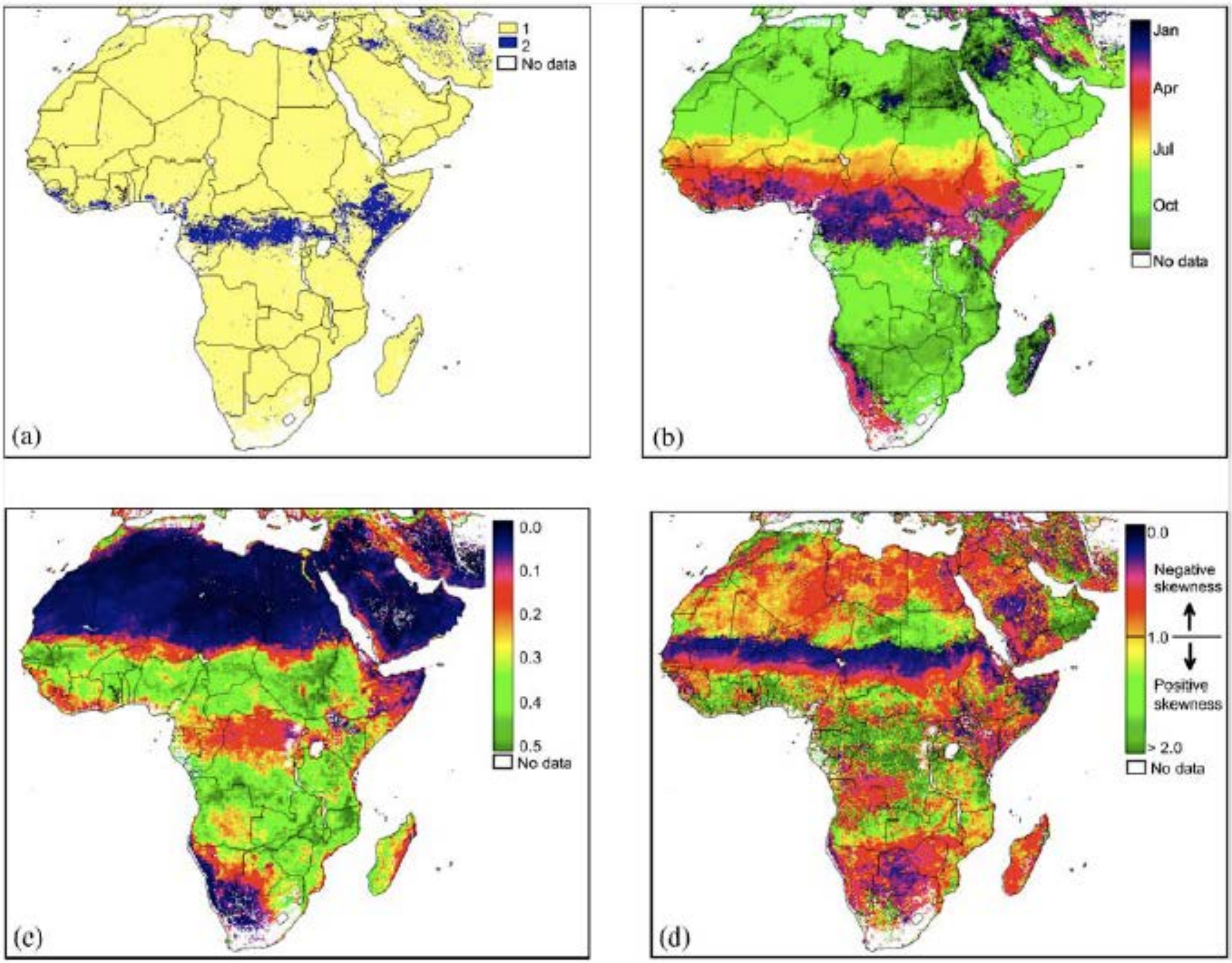
smart filtering example



smart filtering example



smart filtering example



Twicing algorithm

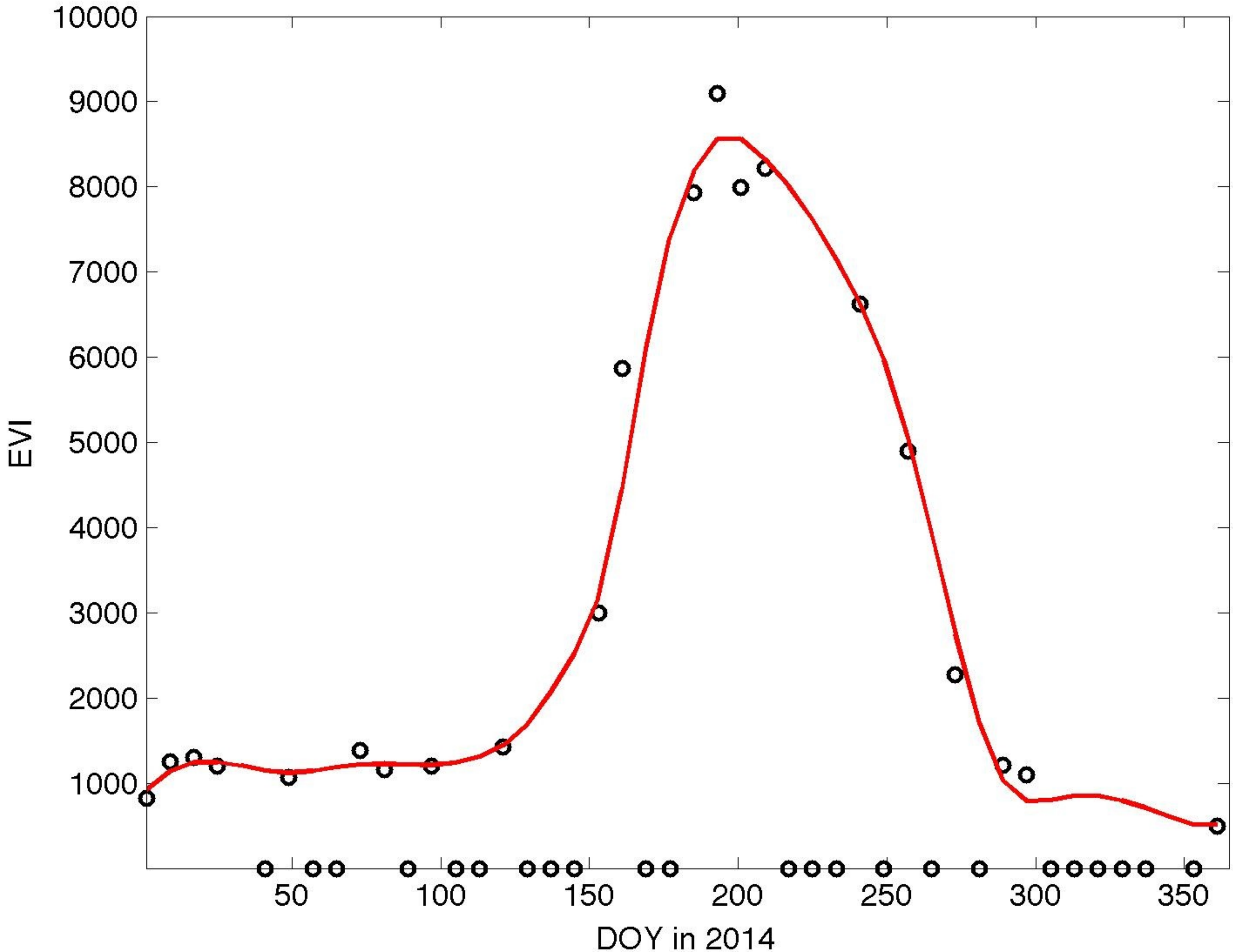
- In the 1980s, a new smoothing algorithm was developed for financial time series smoothing
- It is called twicing because it is a multiple-pass filtering approach

Velleman, P.F., 1980. Definition and Comparison of Robust Nonlinear Data Smoothing Algorithms, Journal of the American Statistical Association, Vol. 75, No. 371 (Sep., 1980), pp. 609-615

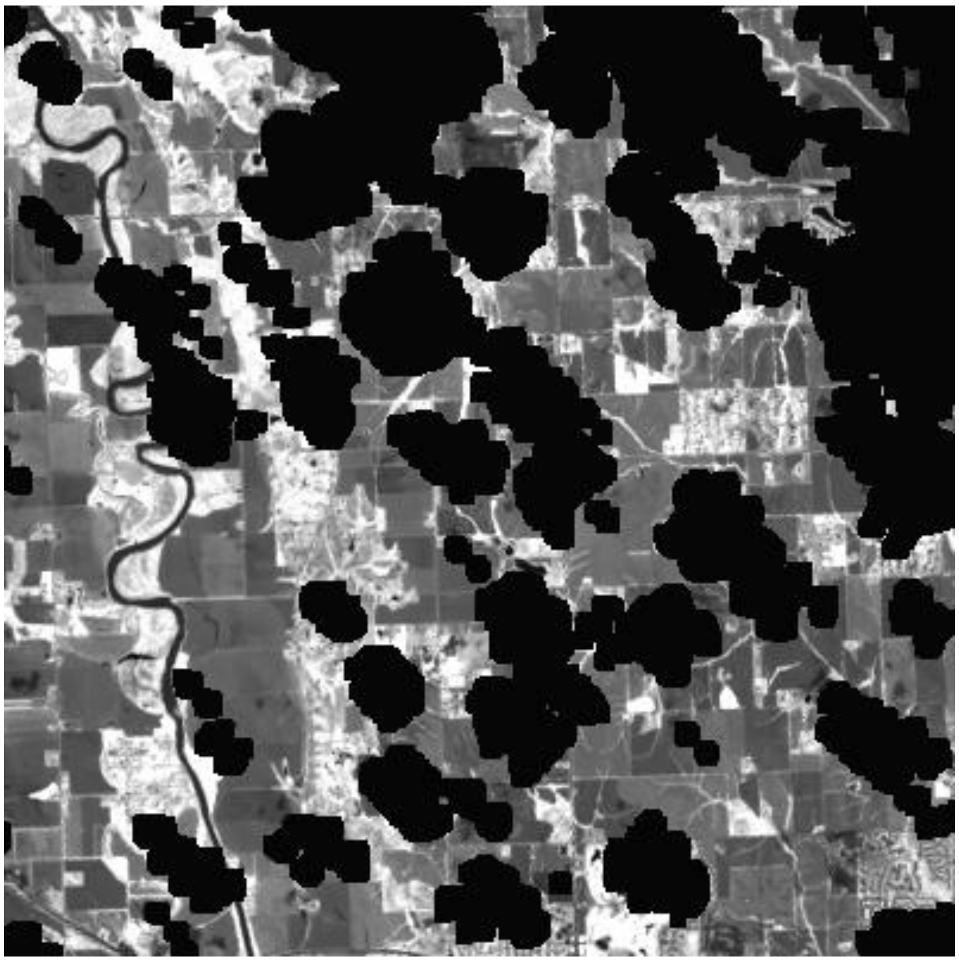
Twicing algorithm

- The algorithm involves taking medians of 4, then 2, then 5, then 3, then Hanning and then by "re-roughing" i.e. by applying the same median smoothing to the residuals of the first pass and adding this to the first pass smoother. Endpoints are always handled using smoothers of shorter, even span or odd span. Hanning is a running weighted average, the weights being 1/4, 1/2 and 1/4.

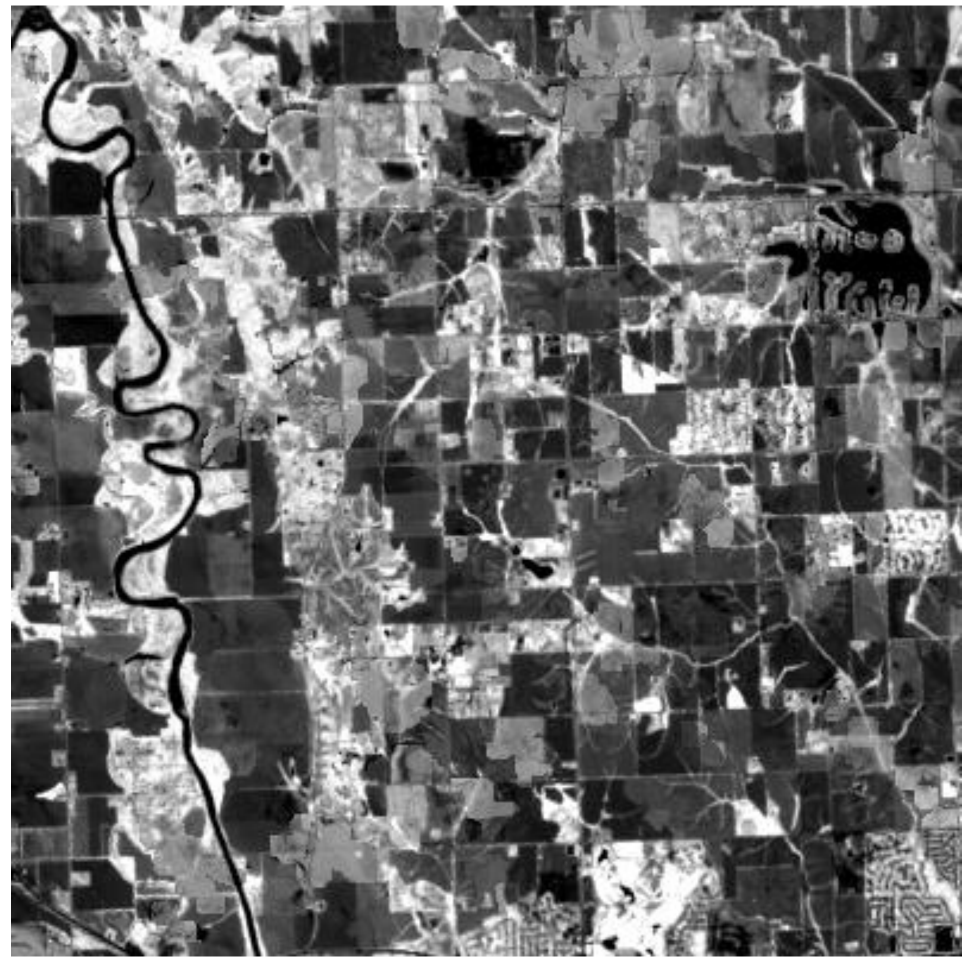
original and smoothed data

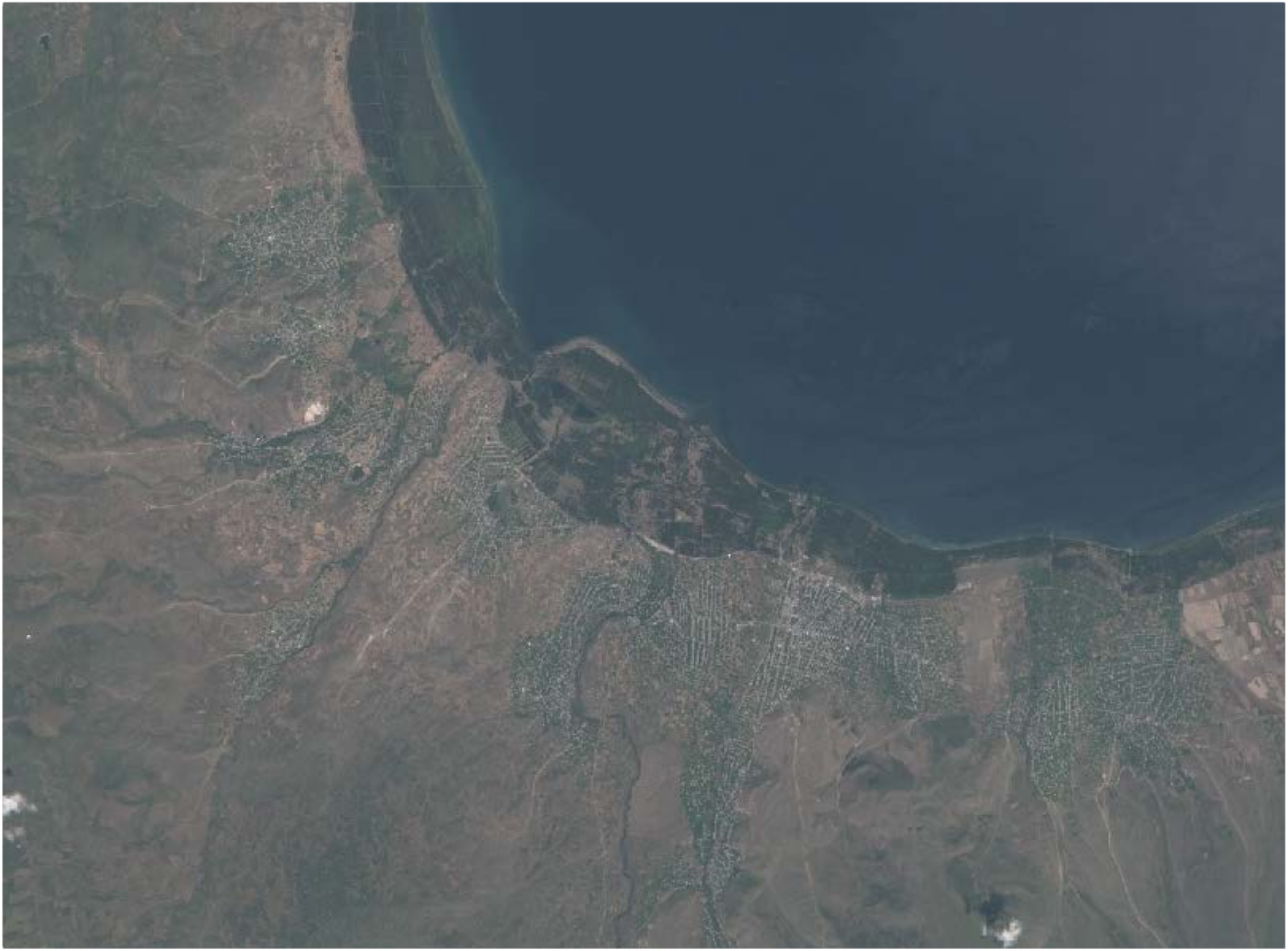


original Landsat EVI



gapfilled Landsat EVI

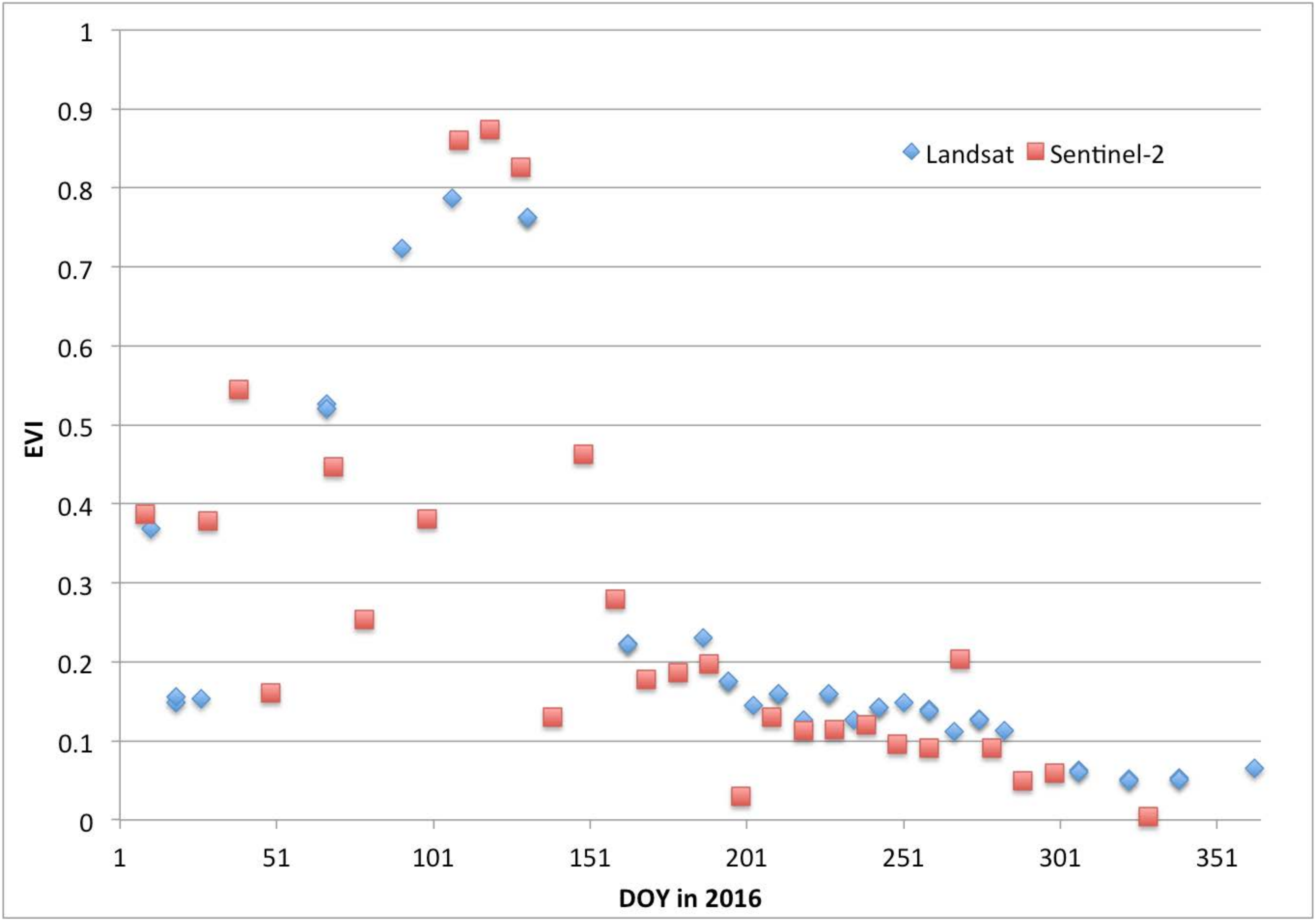




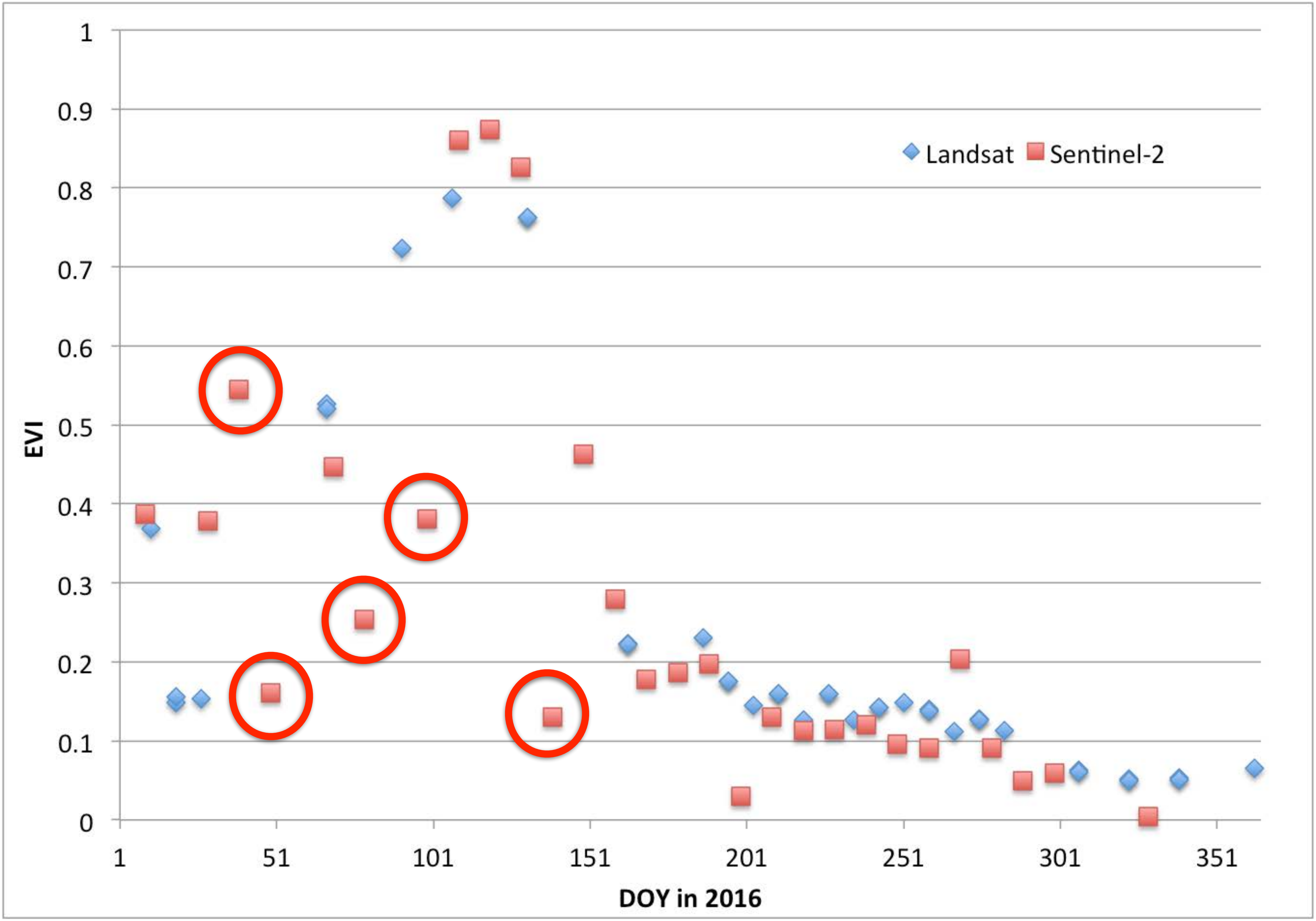
Simple Landsat - Sentinel harmonization

- For agricultural application, data density and continuity is very important
- We developed a “simple” empirical harmonization between Landsat and Sentinel 2A based on regression in the vegetation index space
- Apply regression to “correct” Sentinel-2A

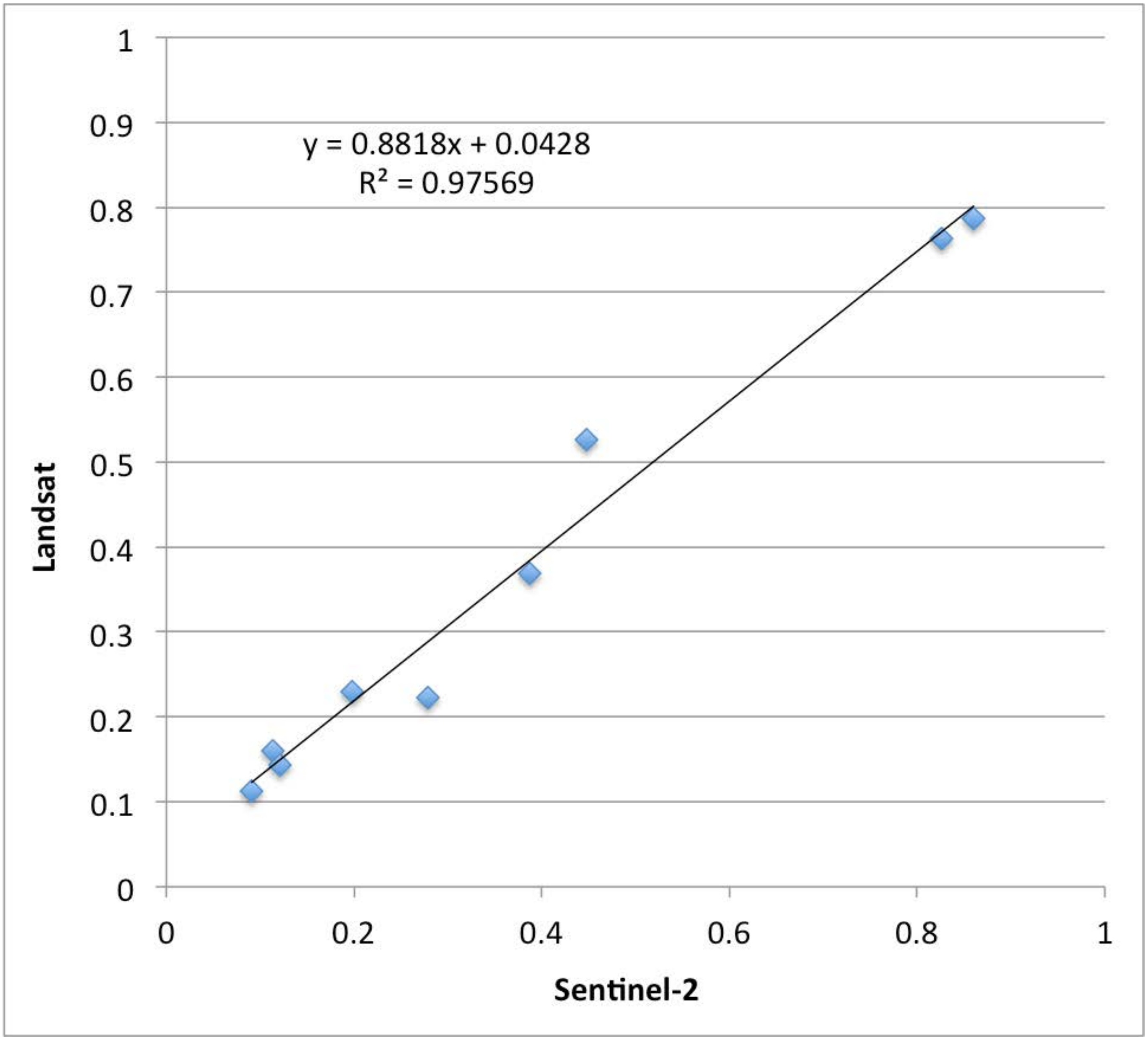
A winter wheat field in Turkey



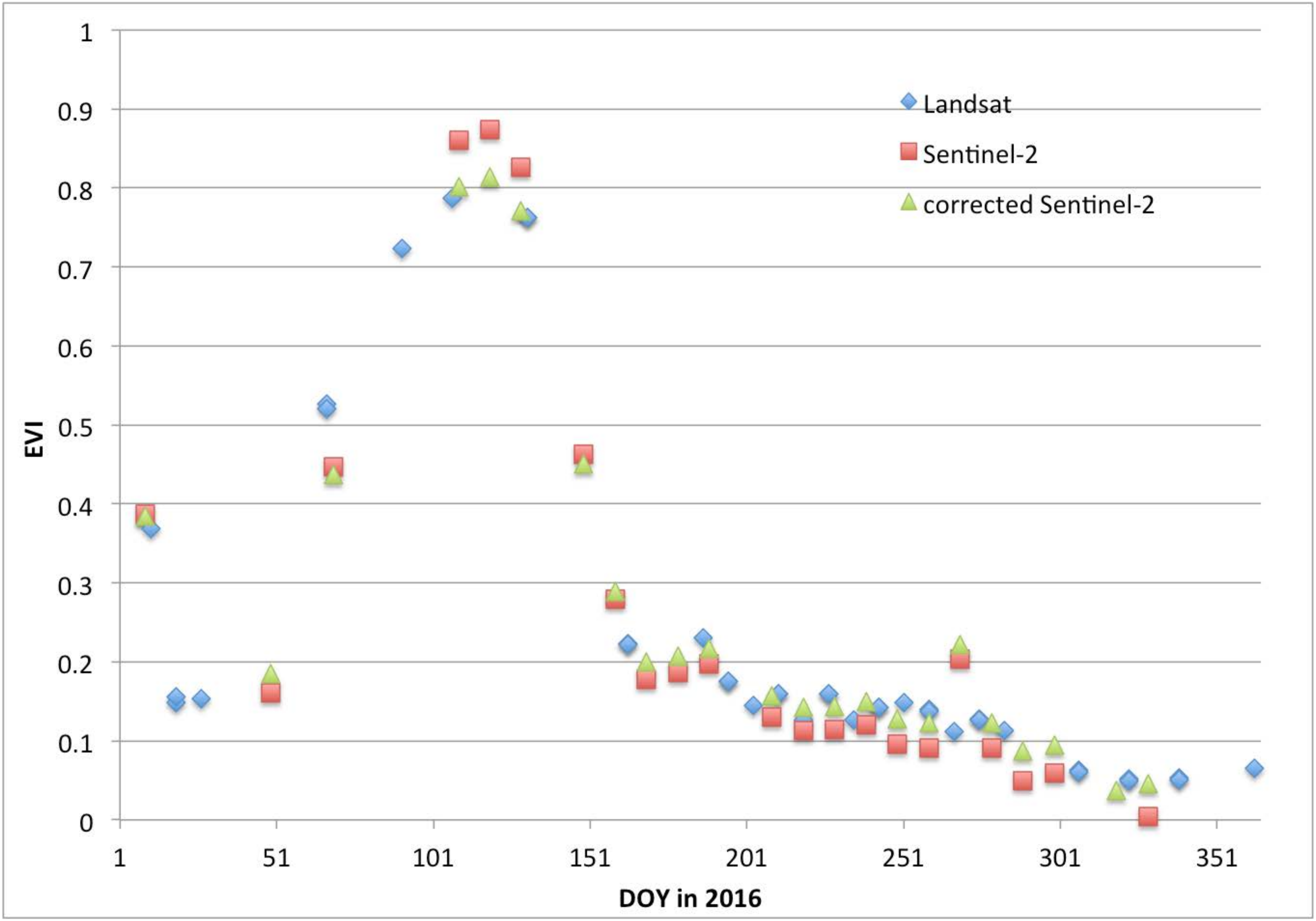
A winter wheat field in Turkey



relationship between Landsat and Sentinel-2A EVI (+- 5days from each other)

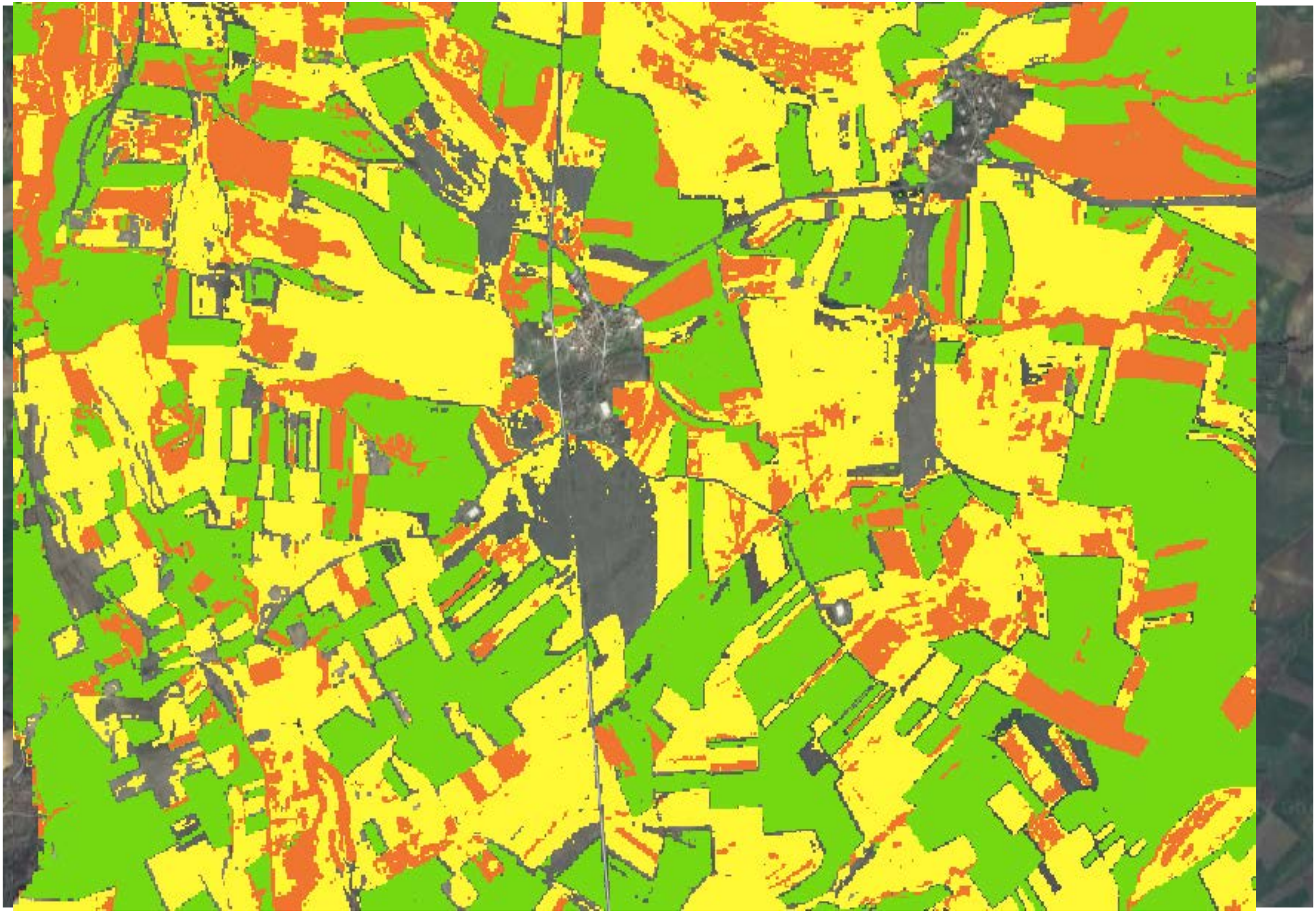


A winter wheat field in Turkey



crop classification in Turkey

crop classification in Turkey



Time series analysis methods

Pre-processing

- gap fill
- averaging (smoothing)
- smart filtering

Information extraction

- Classification
- Phenological metrics
- Fourier analysis
- temporal unmixing
- many more....

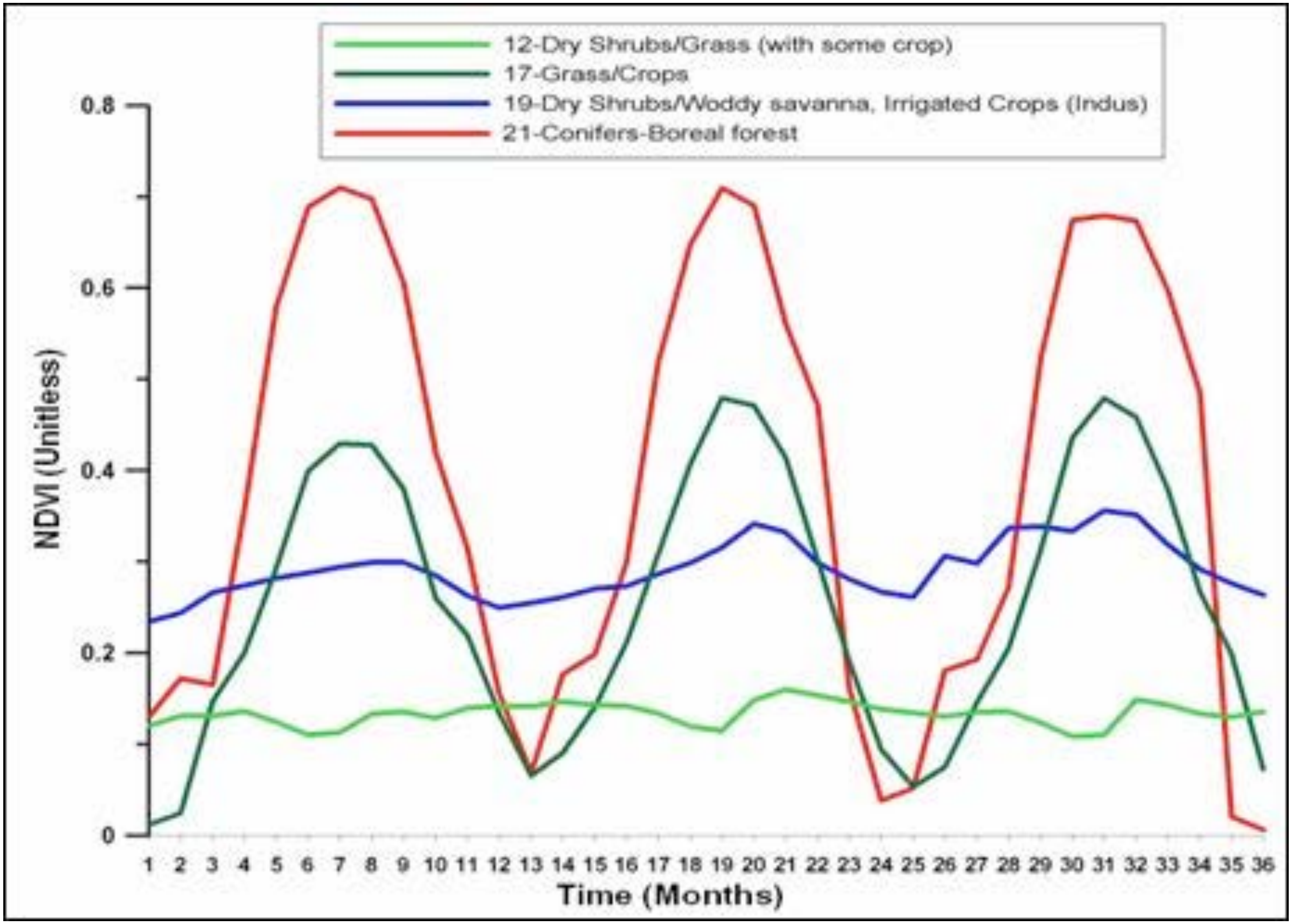
Information extraction

time series classification

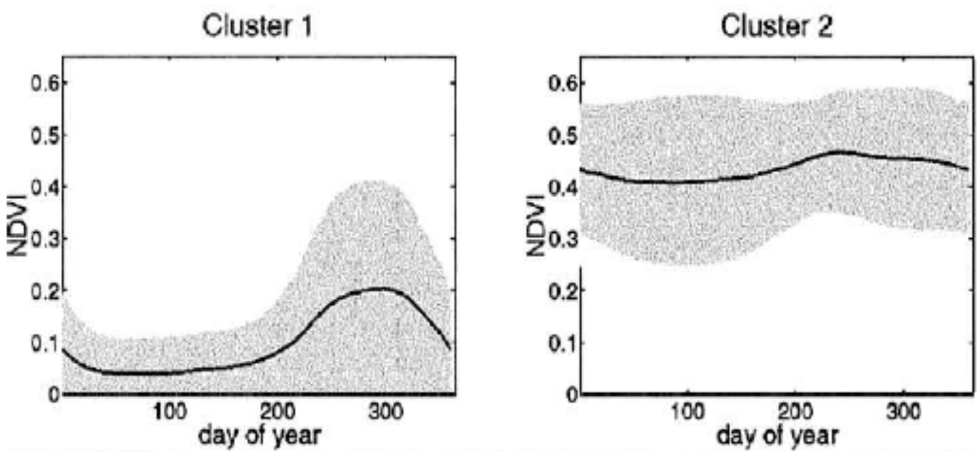
- Classification of time series data: Similar to spectral classification, time-series classification takes the time series data as the input and uses either supervised or unsupervised classification methods to label individual pixels
- Since the main input is a time series data, the assumption here is that the land cover classes of interest are distinguishable partly because of their temporal characteristics and not because of their spectral characteristics
- Uses same tools available, but there are issues related to normality and missing data

a clustering example

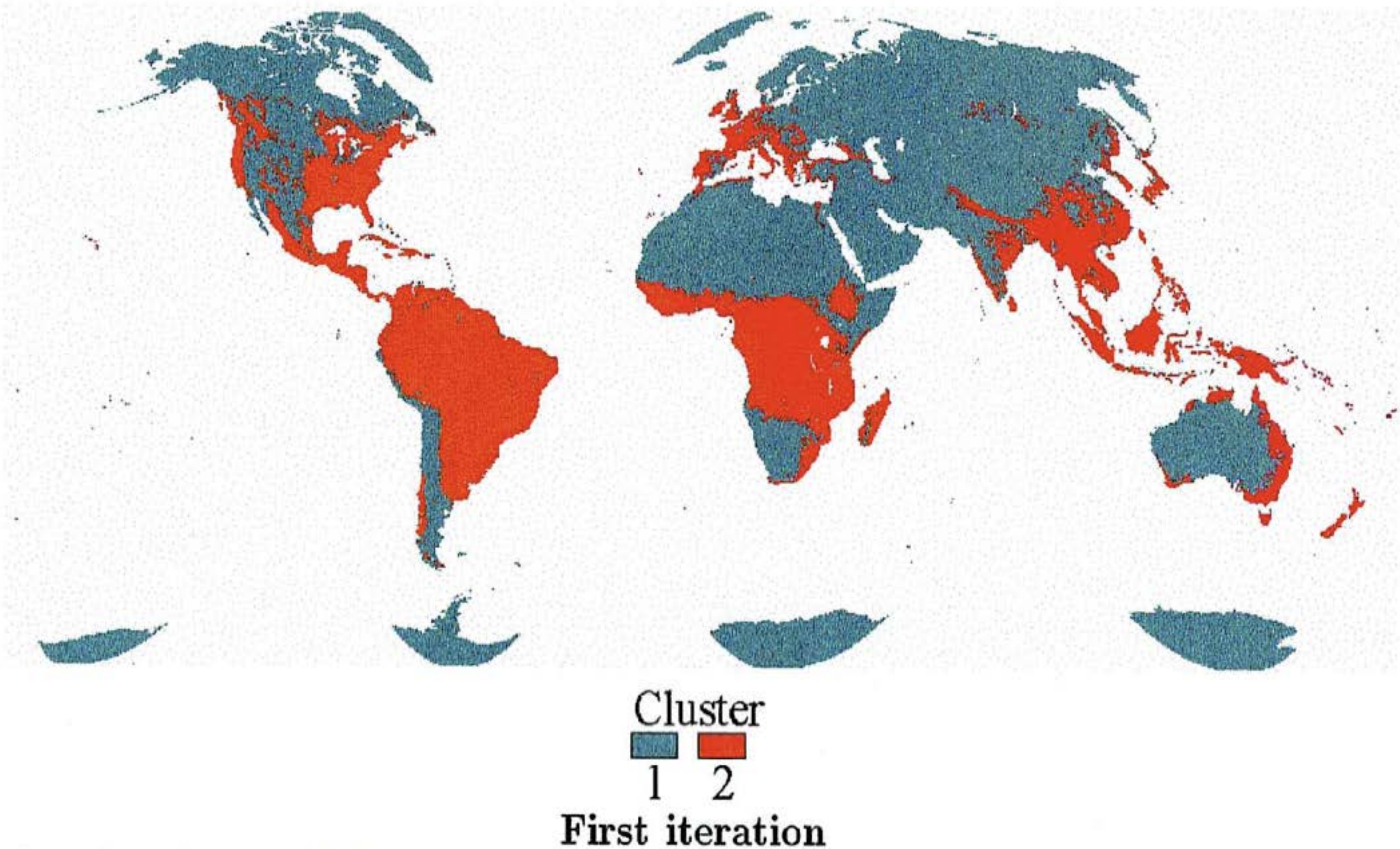
temporal separability



First iteration

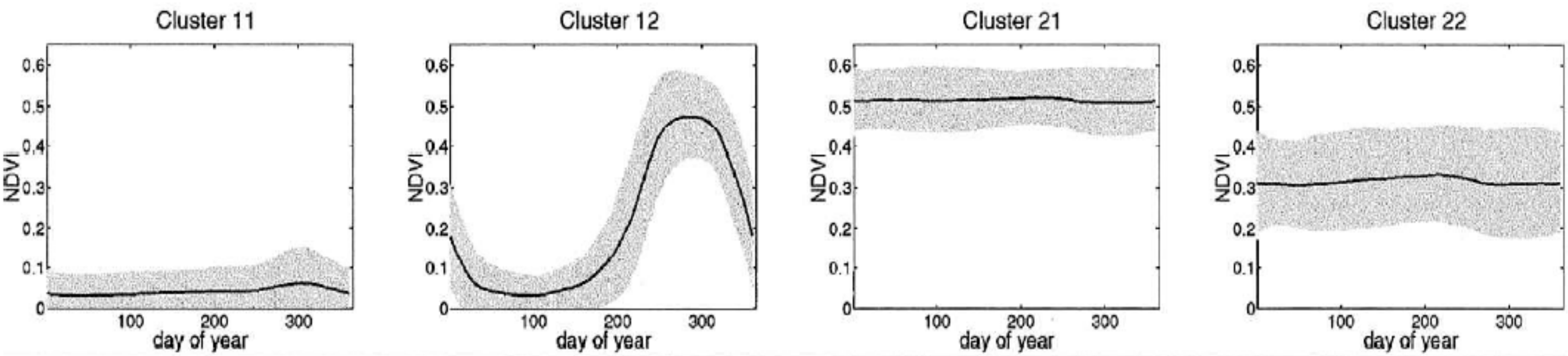


Viovy (2000)

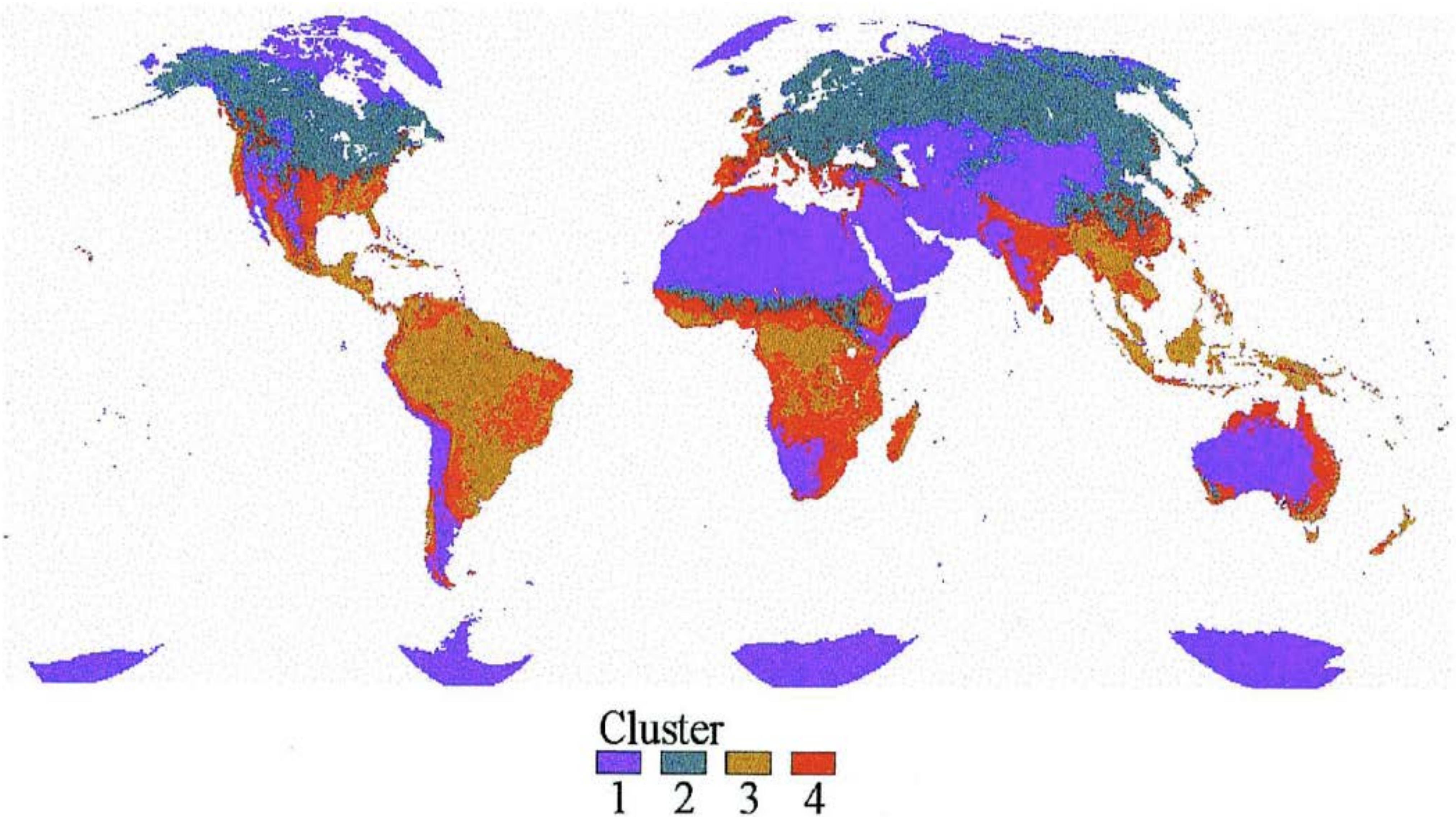


Viovy (2000)

Second iteration



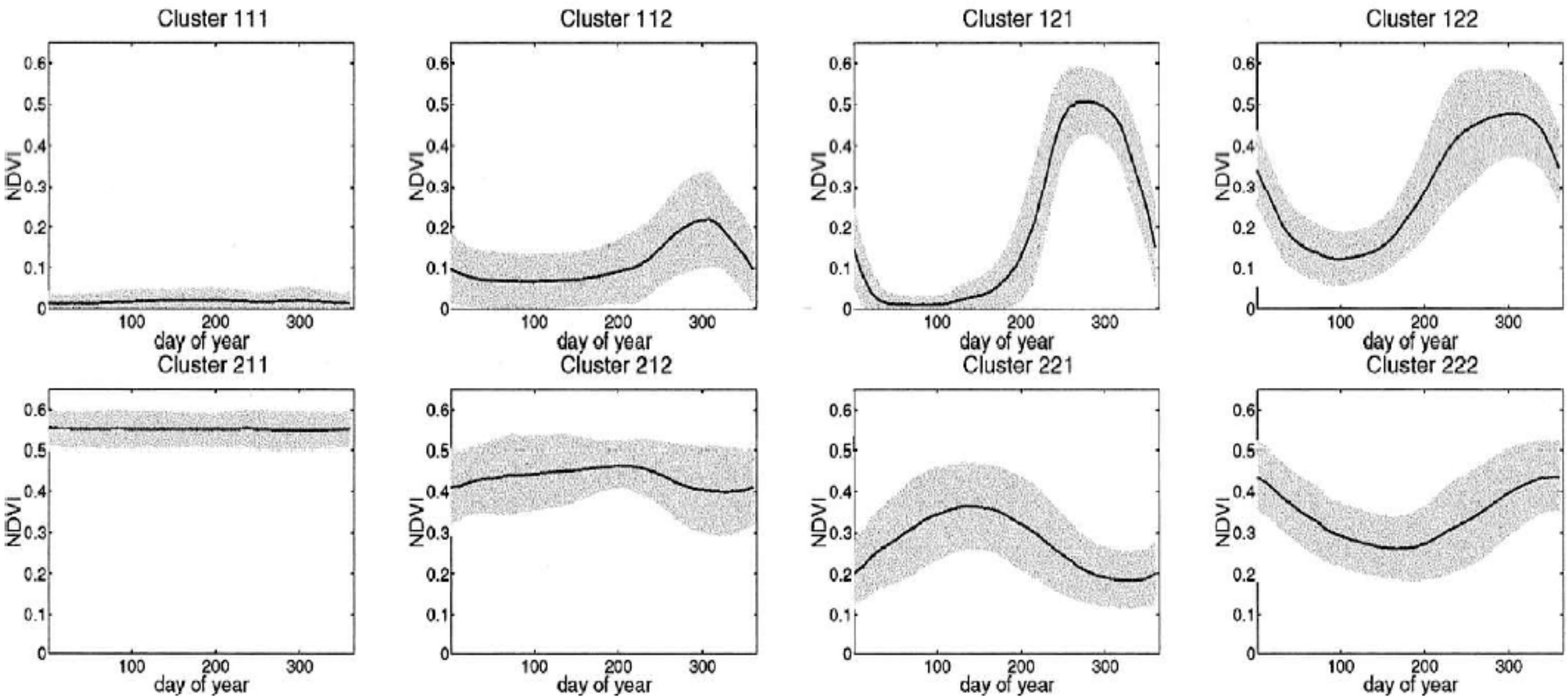
Viovy (2000)



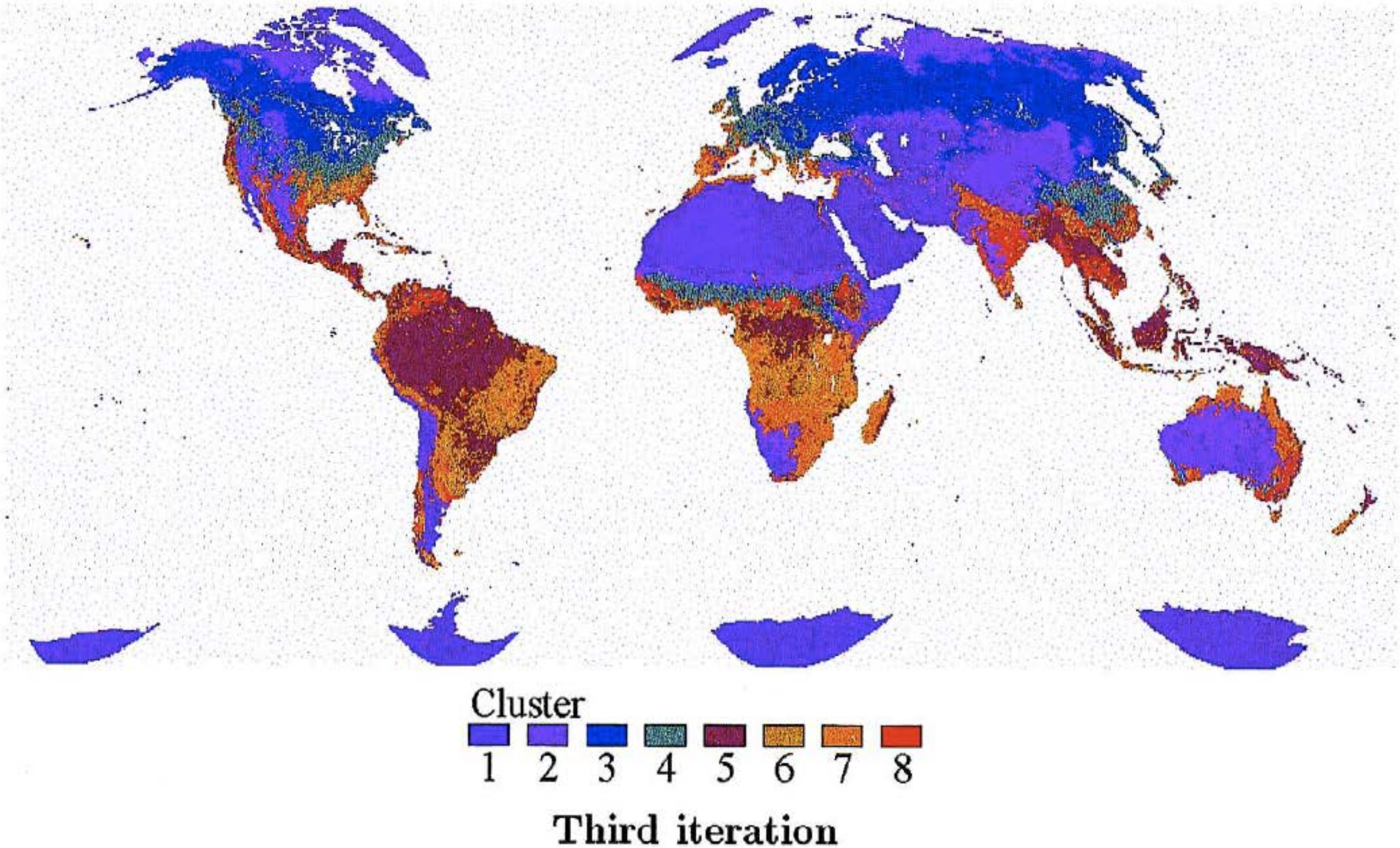
Second iteration

Viovy (2000)

Third iteration

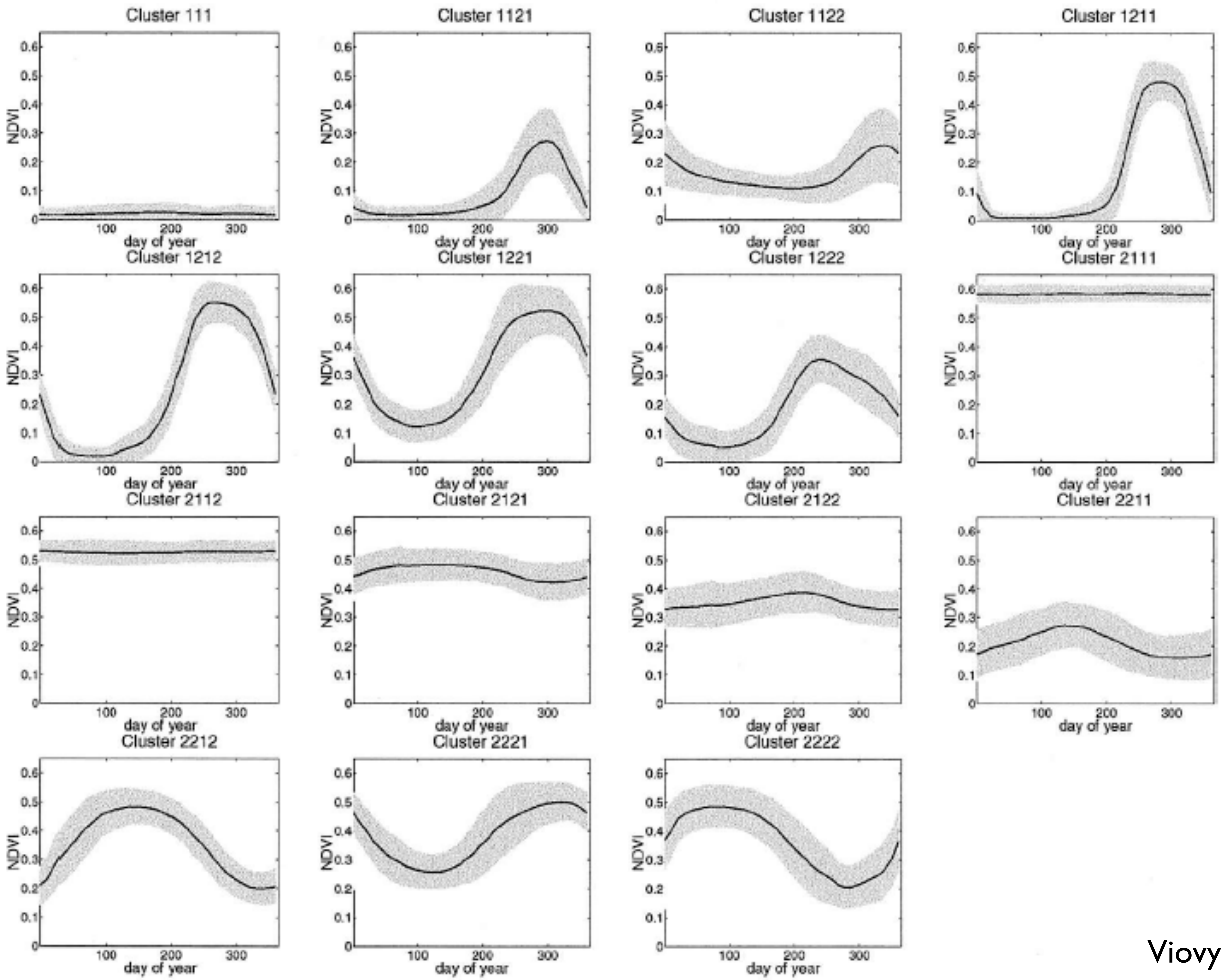


Viovy (2000)

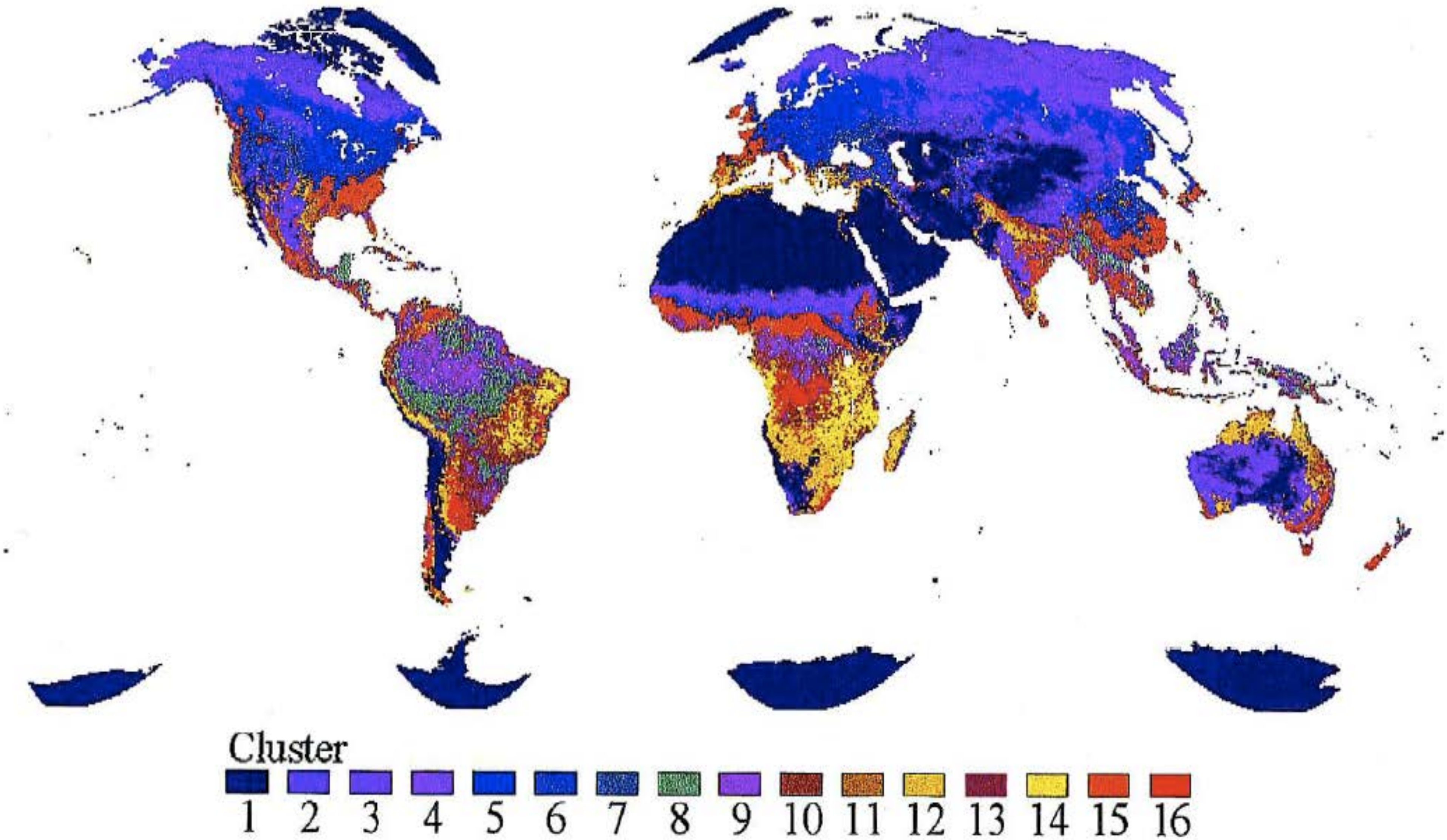


Viovy (2000)

Fourth iteration



Viovy (2000)



Fourth iteration

Viovy (2000)

Fourier Analysis (in the temporal domain)

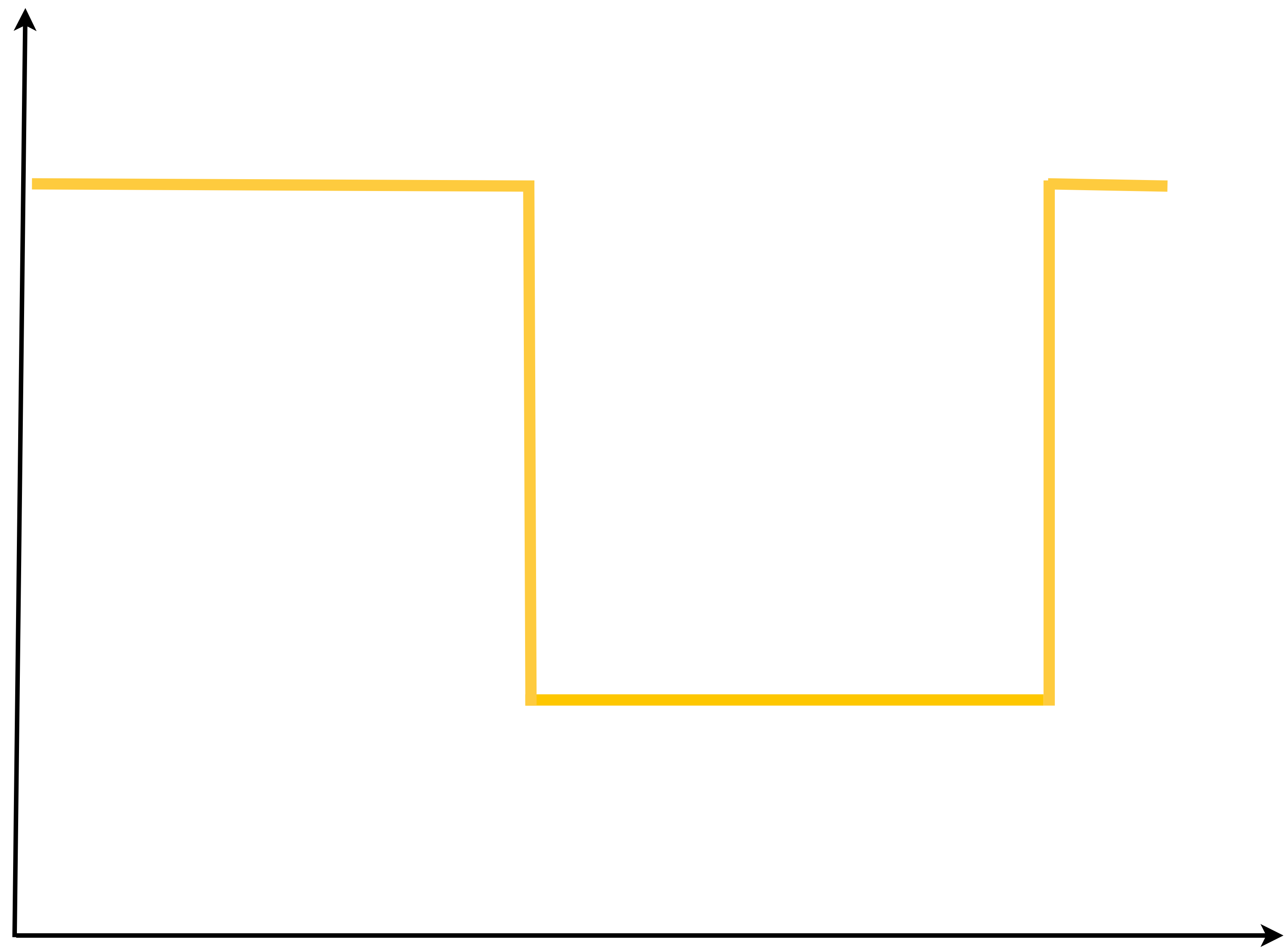
(named after Joseph Fourier who showed that representing a function by a trigonometric series greatly simplified the study of heat propagation)

What is Fourier Analysis?

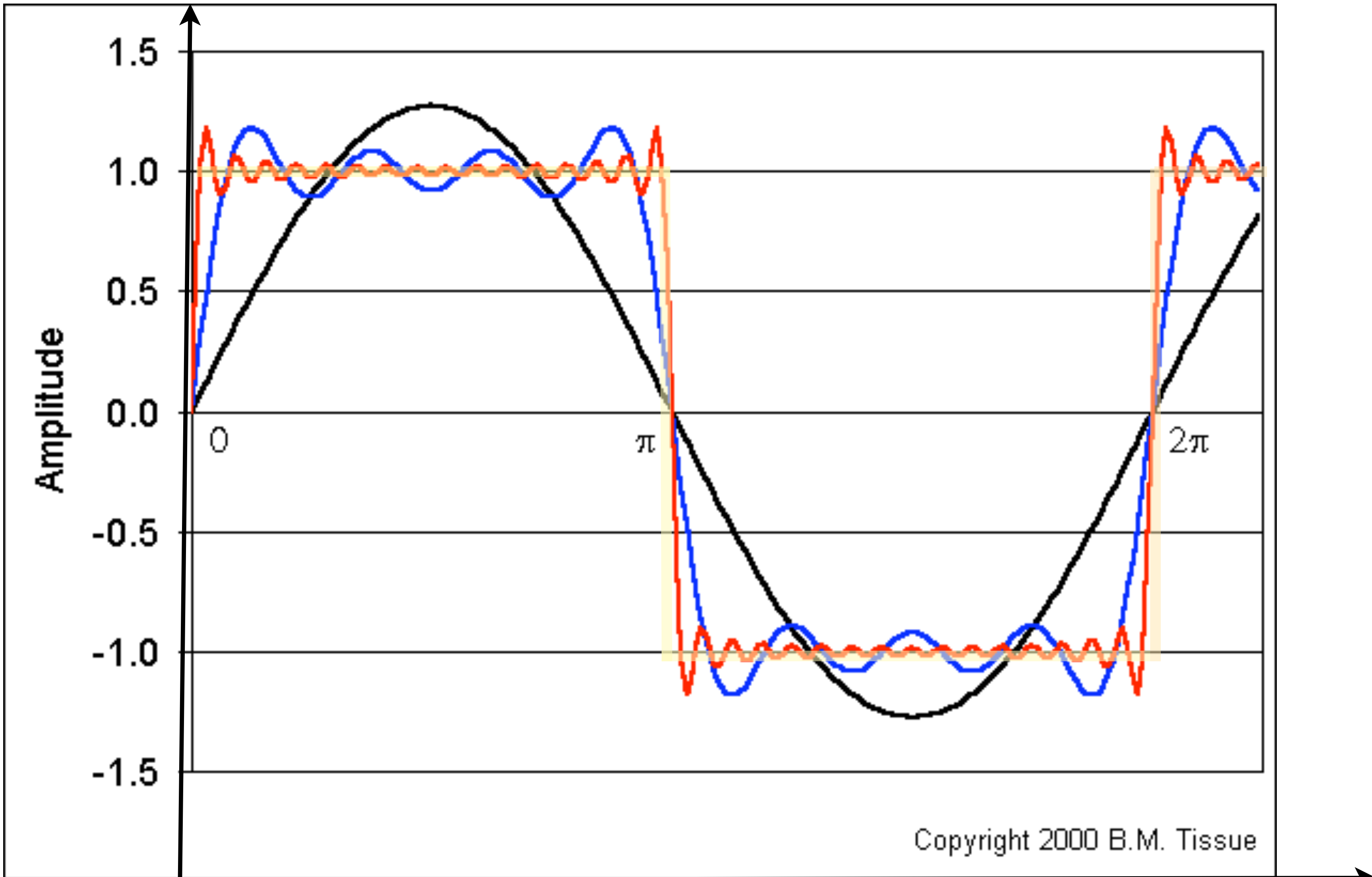
- In its most basic form, Fourier Analysis is an analytical process to decompose a given complex function into simpler (often trigonometric) parts (functions)

- Fourier analysis describes the process of breaking the function into a sum of simpler pieces while Fourier synthesis describes the act of rebuilding the signal

Approximations of a square wave



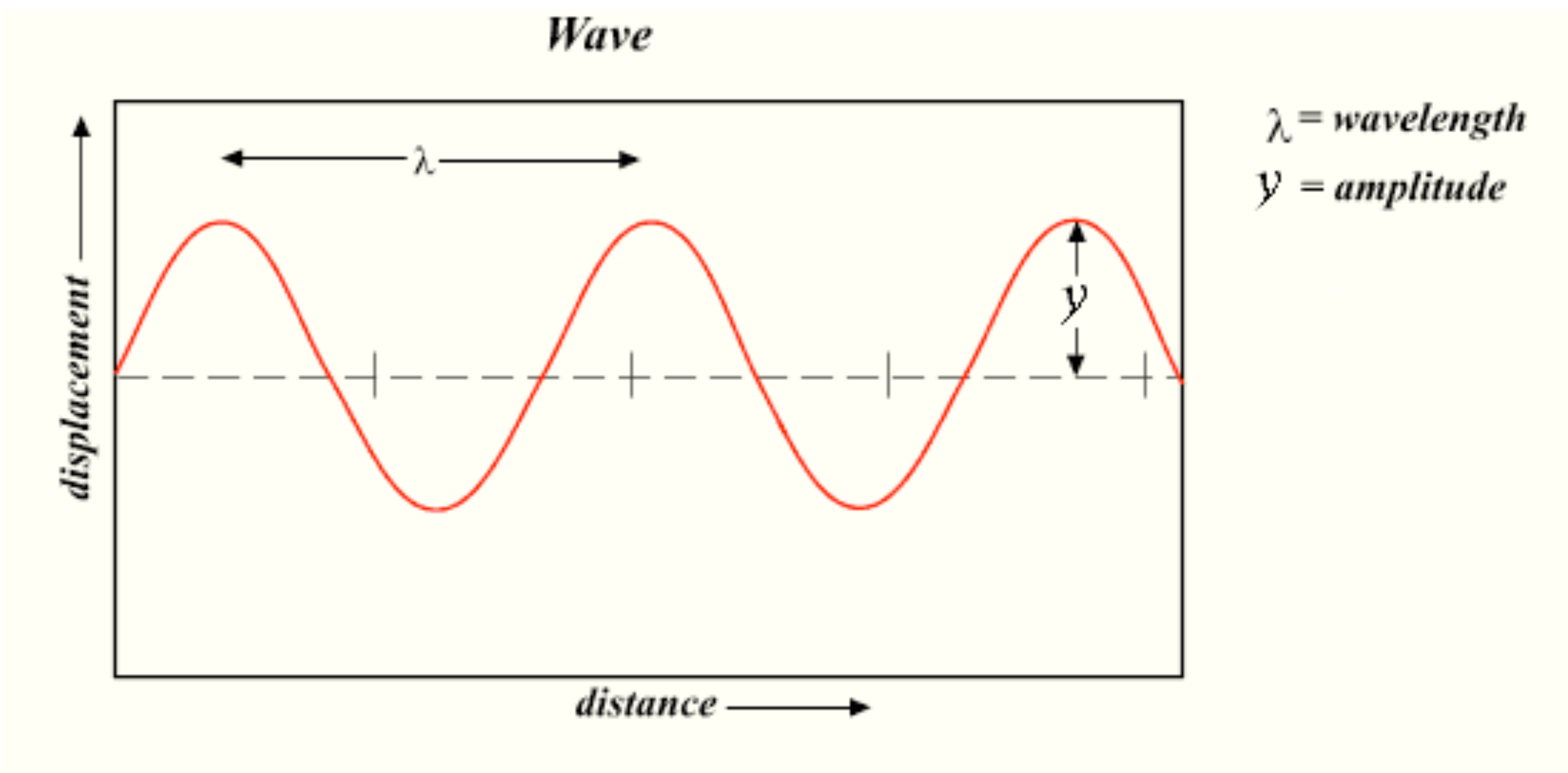
Approximations of a square wave



Black - 1 term approximation
Blue - 4 term approximation
Red - 16 term approximation

Some basic terms

- Wavelength: the distance over which the wave's shape repeats
- Frequency: the number of occurrences of a repeating event per unit time
- Period: the duration of one cycle in a repeating event, so the period is the reciprocal of the frequency



What is Fourier Analysis?

- The Fourier theorem states that any waveform can be duplicated by the superposition of a series of sine and cosine waves.
- The Fourier Transform is used convert between two descriptions of a physical system:

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)\, e^{-i\omega t}\, d\omega$$

$$F(\omega) = \int_{-\infty}^{\infty} f(t)\, e^{i\omega t}\, dt$$

- The F(W) function gives the frequencies at which the signal is non-zero and the f(t) function gives the times at which the signal is non-zero. Both of these functions are suitable descriptions of a waveform or physical system.

Discrete Fourier Transform

- in digital images, we often deal with Discrete Fourier Transform (DFT) which contains only the sampled frequencies, and not all of the frequencies forming an image but enough samples to fully describe an image.
- For a square image with *N* x *N* size, DFT is given by:

$$F(k,l) = \frac{1}{N^2} \sum_{a=0}^{N-1} \sum_{b=0}^{N-1} f(a,b)\, e^{-\iota 2\pi \left(\frac{ka}{N} + \frac{lb}{N}\right)}$$

- where *f(a,b)* is the original spatial domain image
- where *F(k,l)* is the Fourier domain image

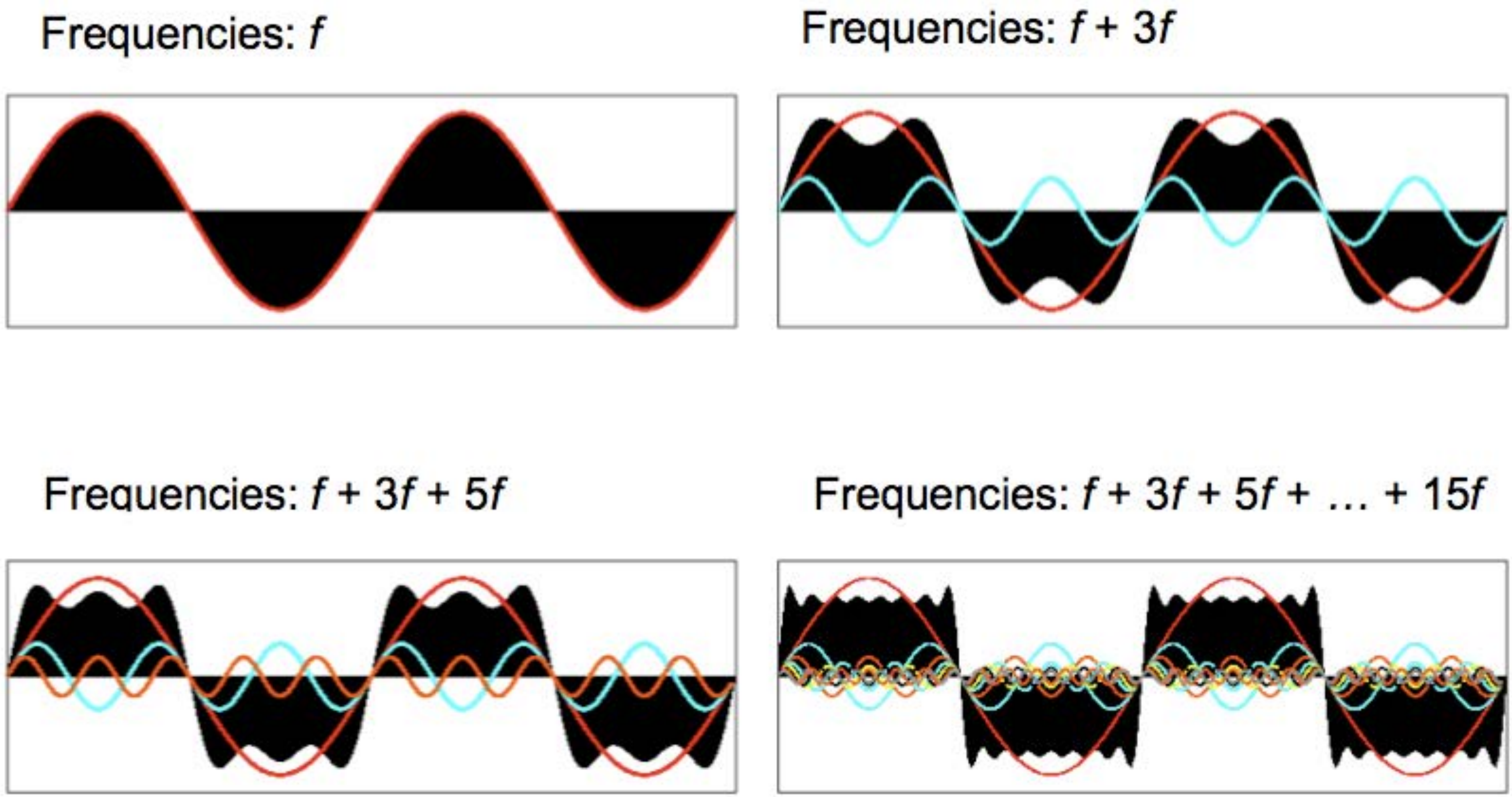
What is Fourier Transform in the temporal domain?

- in temporal analysis, one is often interested in determining the frequency content of temporal signals.
- temporal analysis is based on the notion that a signal in each pixel is characterized by time-dependent functions that are <u>cyclical</u> in nature.
- the purpose of Fourier Transform in the temporal domain is then to decompose a complex time series with cyclical components into a few underlying sinusoidal (sine and cosine) functions of particular wavelengths.

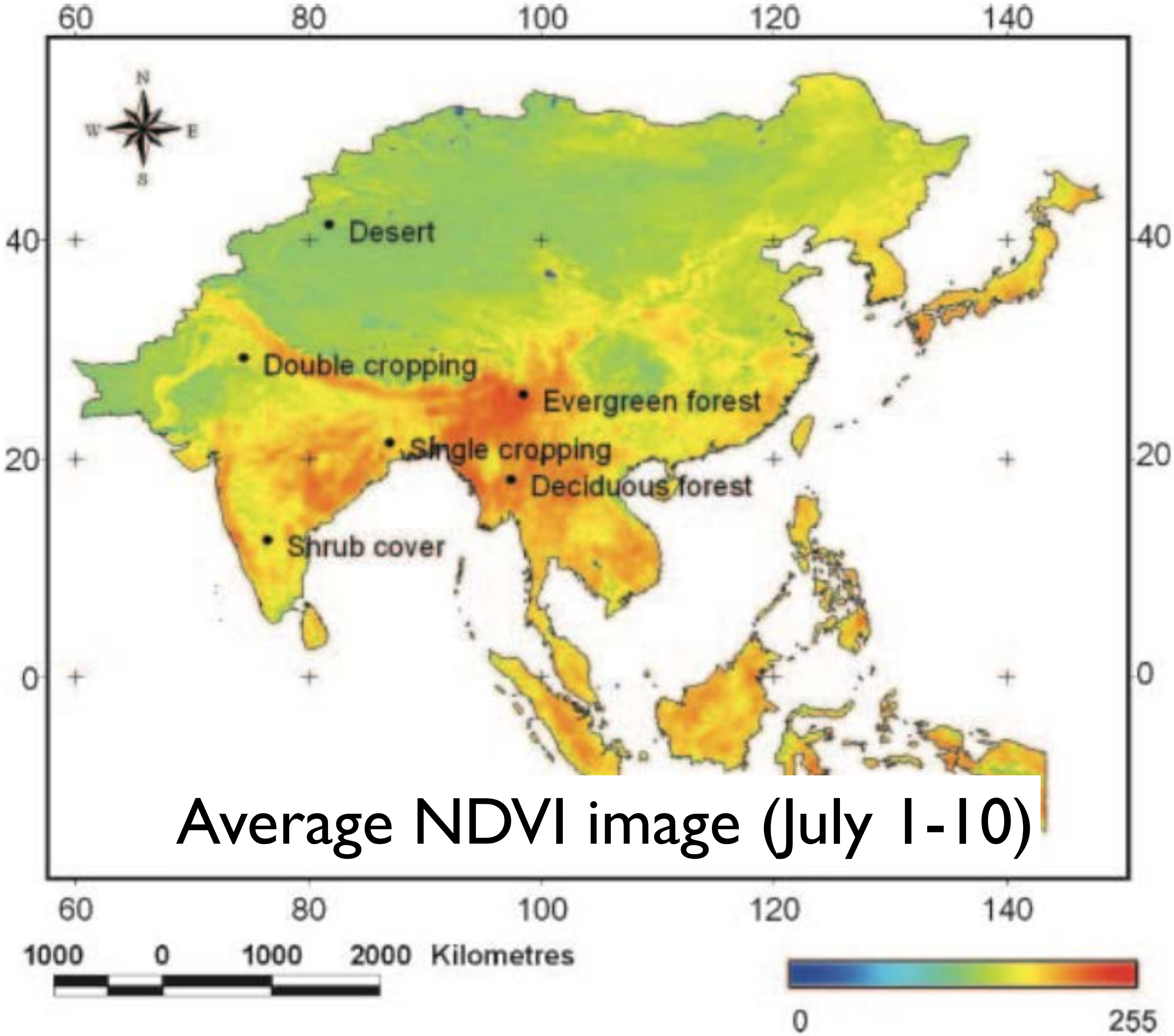
What is Fourier Transform in the temporal domain?

- By identifying the important underlying cyclical components through Fourier Transforms, we can learn something about the environment
- In essence, performing frequency analysis on a time series is like putting the series through a prism in order to identify the wavelengths and importance of underlying cyclical components of a landscape
- As a result of a successful analysis one might uncover just a few recurring cycles of different lengths in the time series of interest, which at first may look more or less like random noise.

What is Fourier Transform? (time domain)

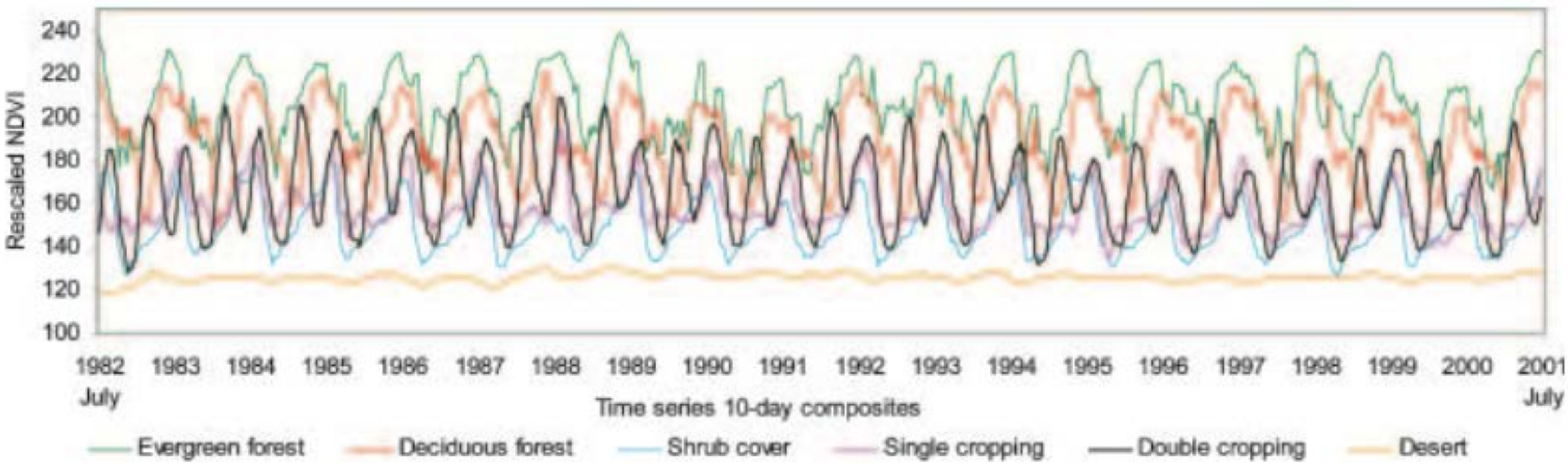


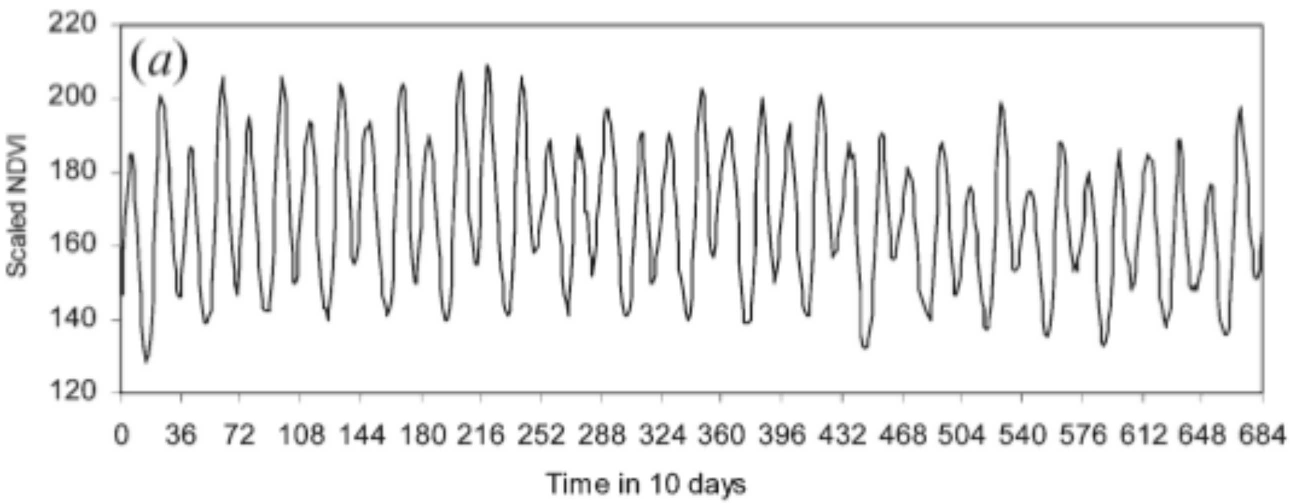
Temporal Fourier Transform Example

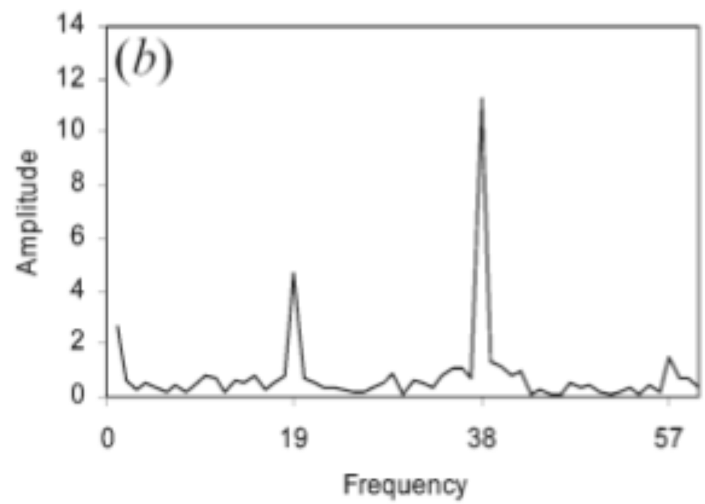


Average NDVI image (July 1-10)

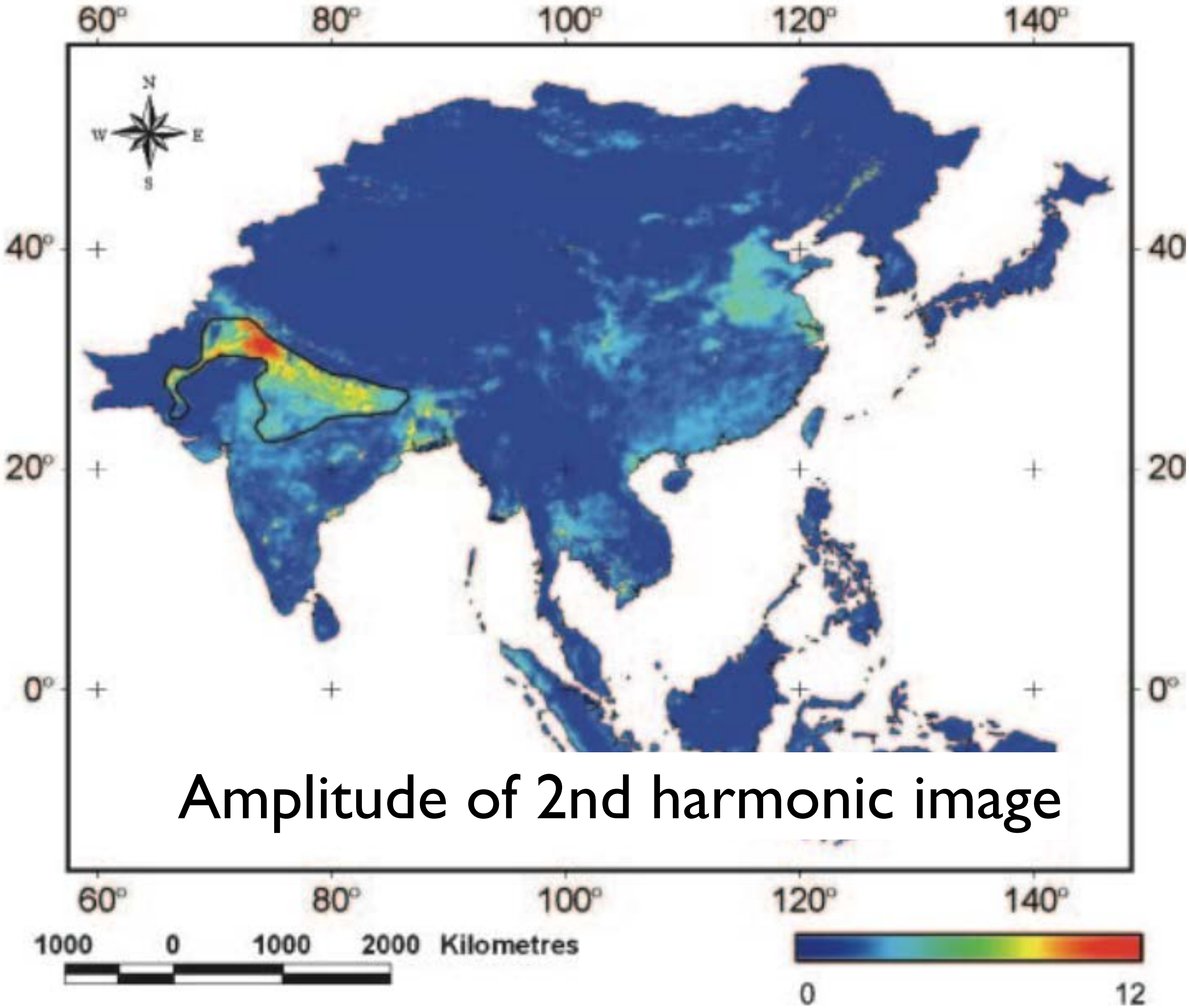
Temporal Fourier Transform Example





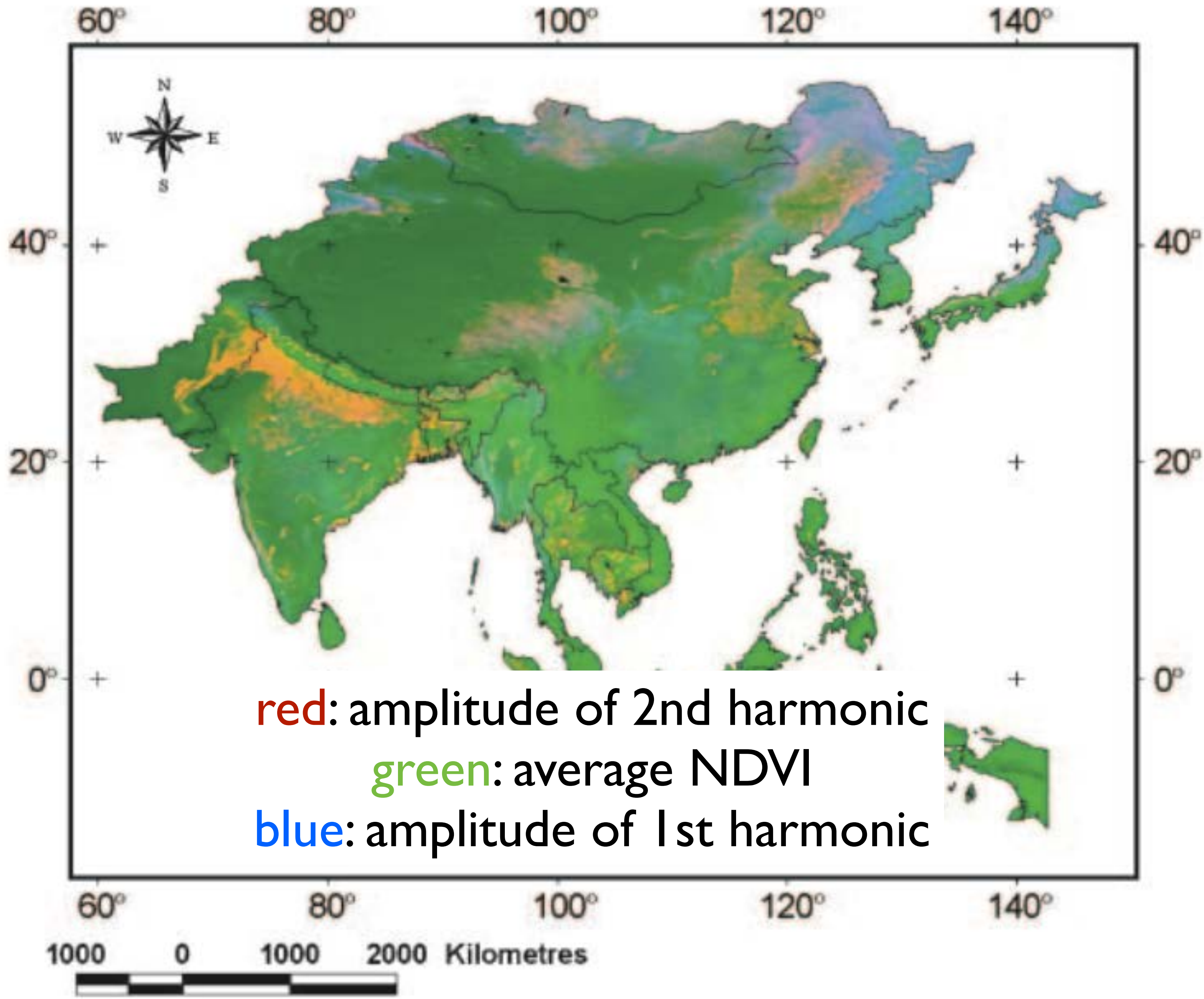


Temporal Fourier Transform Example



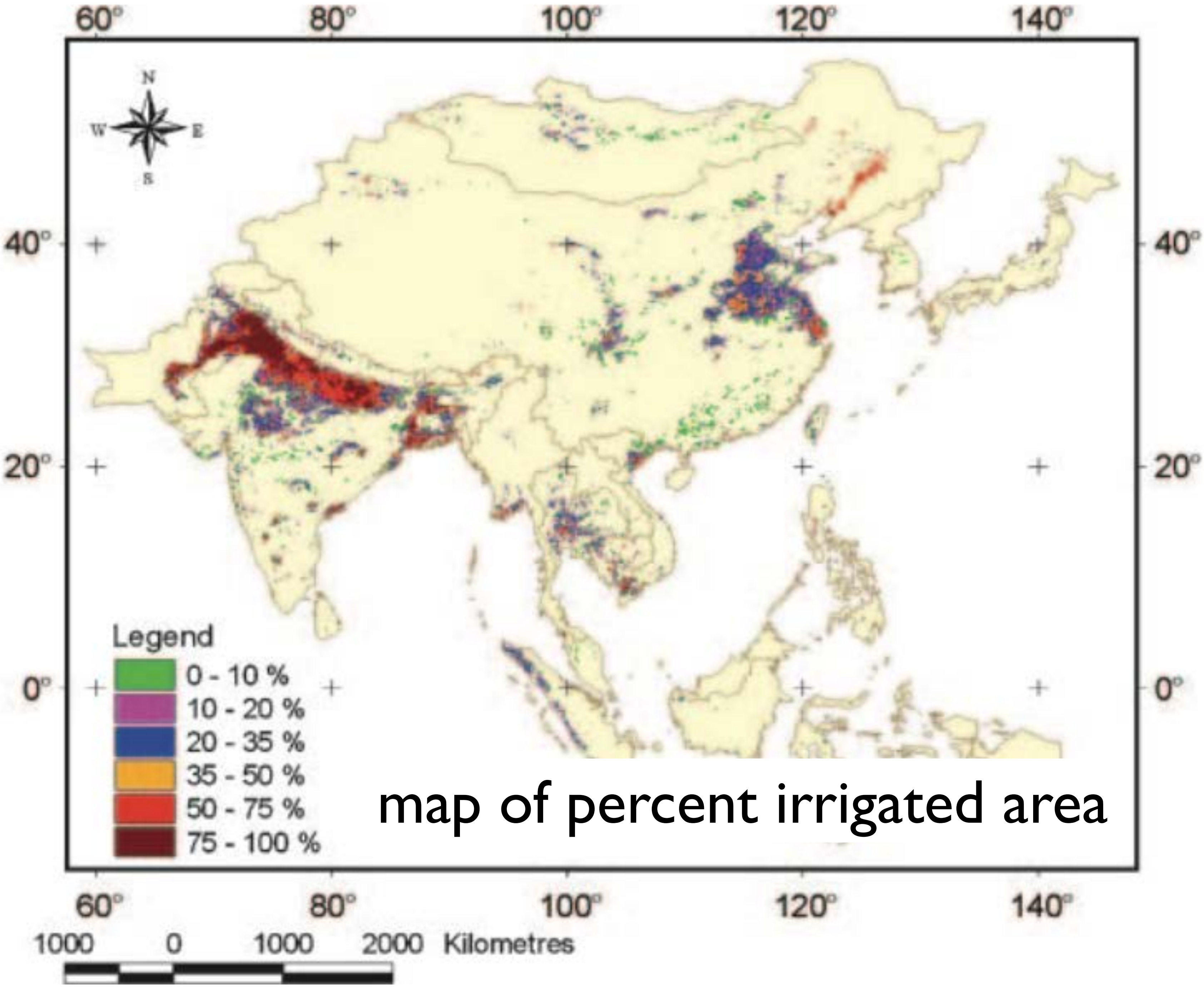
Amplitude of 2nd harmonic image

Temporal Fourier Transform Example



red: amplitude of 2nd harmonic
green: average NDVI
blue: amplitude of 1st harmonic

Temporal Fourier Transform Example



map of percent irrigated area

Dense Time Stacks

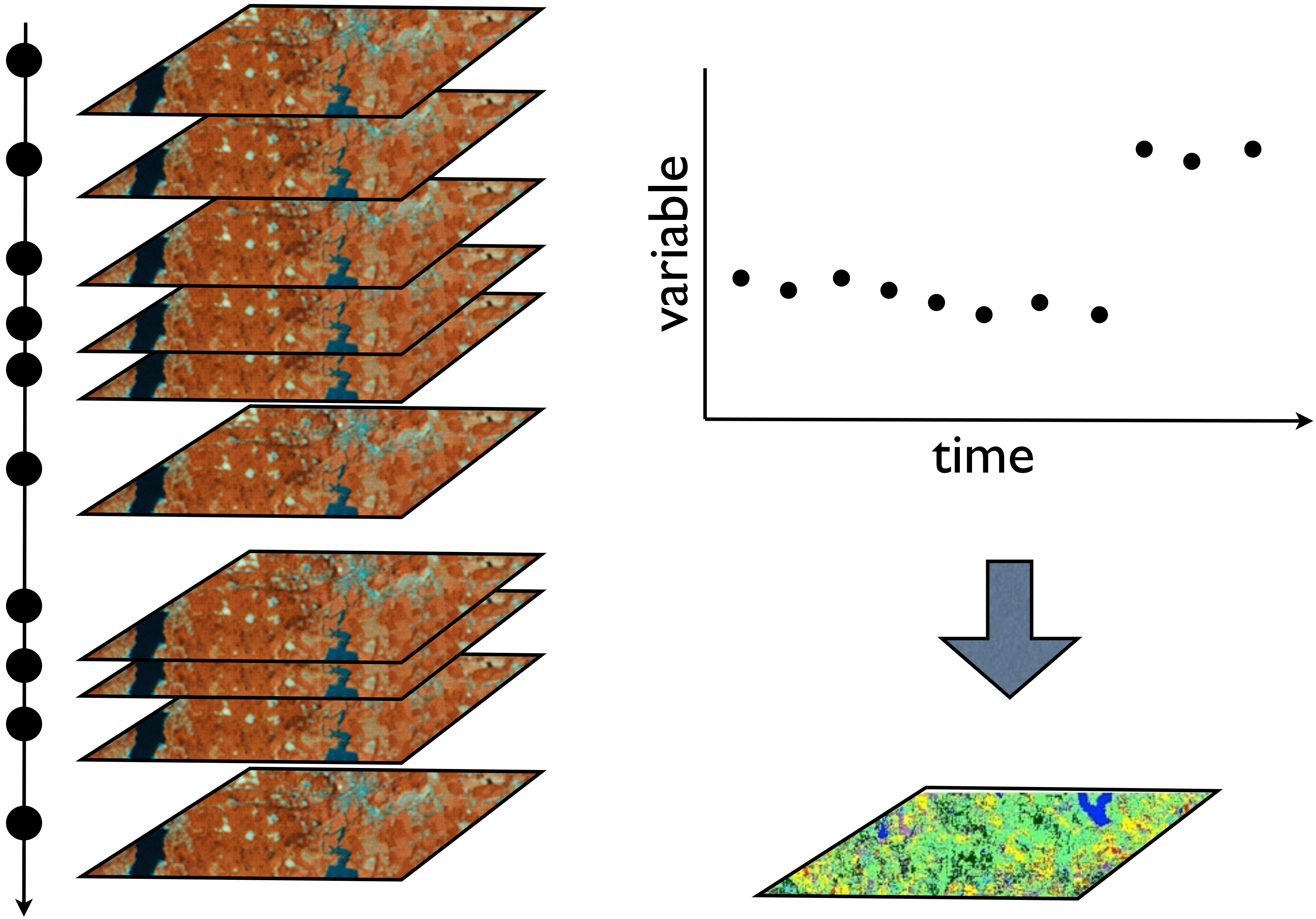
(time series methods applied to Landsat-like data)

What is dense time stacks (DTS)?

- Dense time stacks refers to an image processing concept that involves <u>time compositing</u> of medium spatial resolution (20-100 meter) image data
- The frequency of image composites can range from many per year to one every two years
- This form of image processing of medium resolution data has become possible only recently because of the open and free access to data
- DTS is often use for finding change but not always

What is dense time stacks (DTS)?

- <u>The main difference between DTS based image analysis and Time Series analysis are:</u>
    - Time series analysis often deals with regularly spaced image acquisitions and DTS does not
    - DTS is often used to find change but time series analysis is often used for landscape categorization
    - DTS is used for local to regional applications whereas time series analysis is used for regional to global applications


variable
time

Three methods in the literature

- LandTrendr (Landsat-based Detection of Trends in Disturbance and Recovery)

Kennedy, R.E., Yang, Z., and Cohen, W.B., 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr — Temporal segmentation algorithms, Remote Sensing of Environment, 114:2897-2910.

- VCT (Vegetation Change Tracker)

Huang, C., Goward, S.N., Masek, J.G., Thomas, N., Zhu, Z., and Vogelmann, J.E., 2010. An automated approach for reconstructing recent forest disturbance history usingdense Landsat time series stacks, Remote Sensing of Environment, 114:183-198

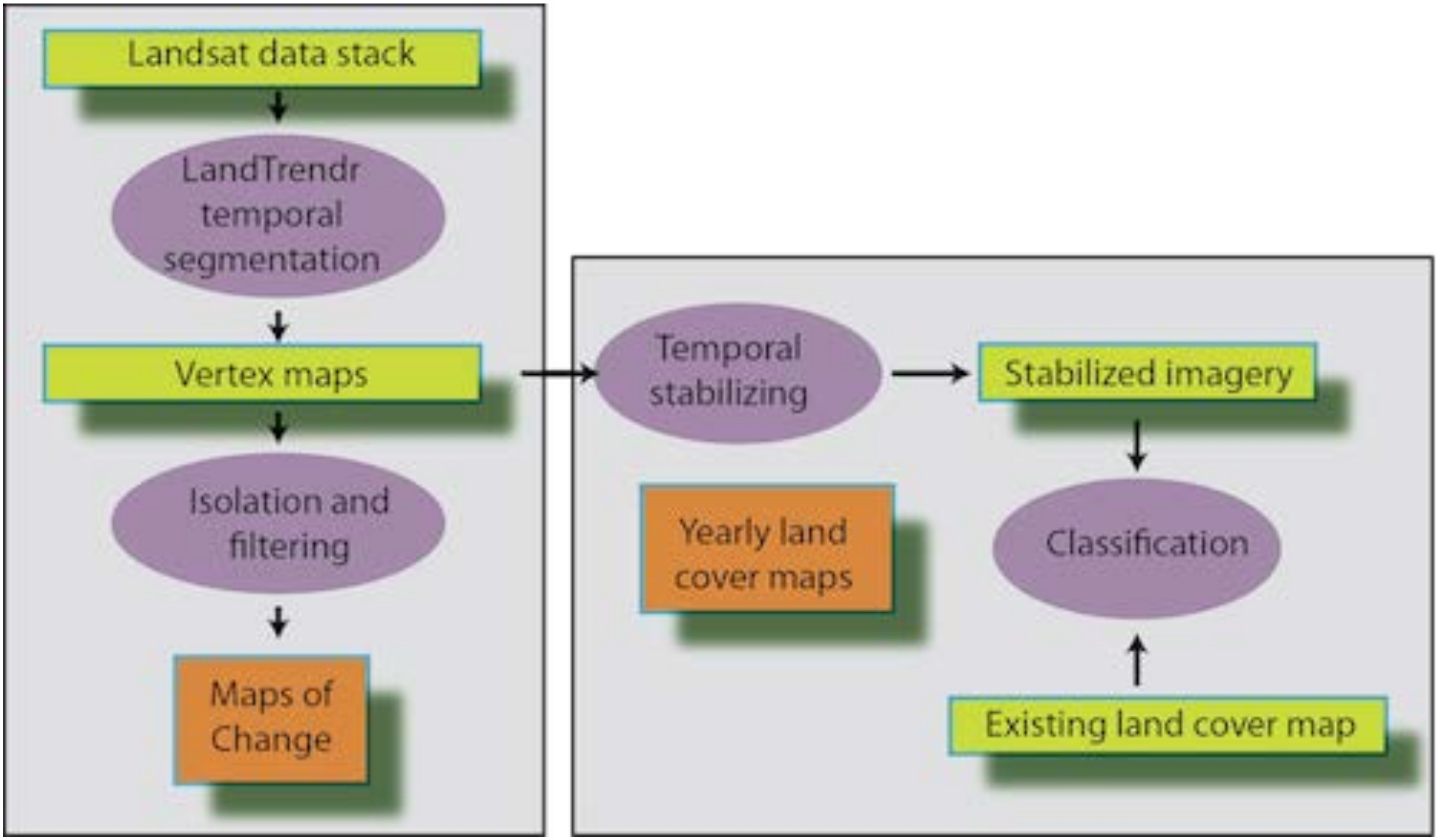
Three methods in the literature

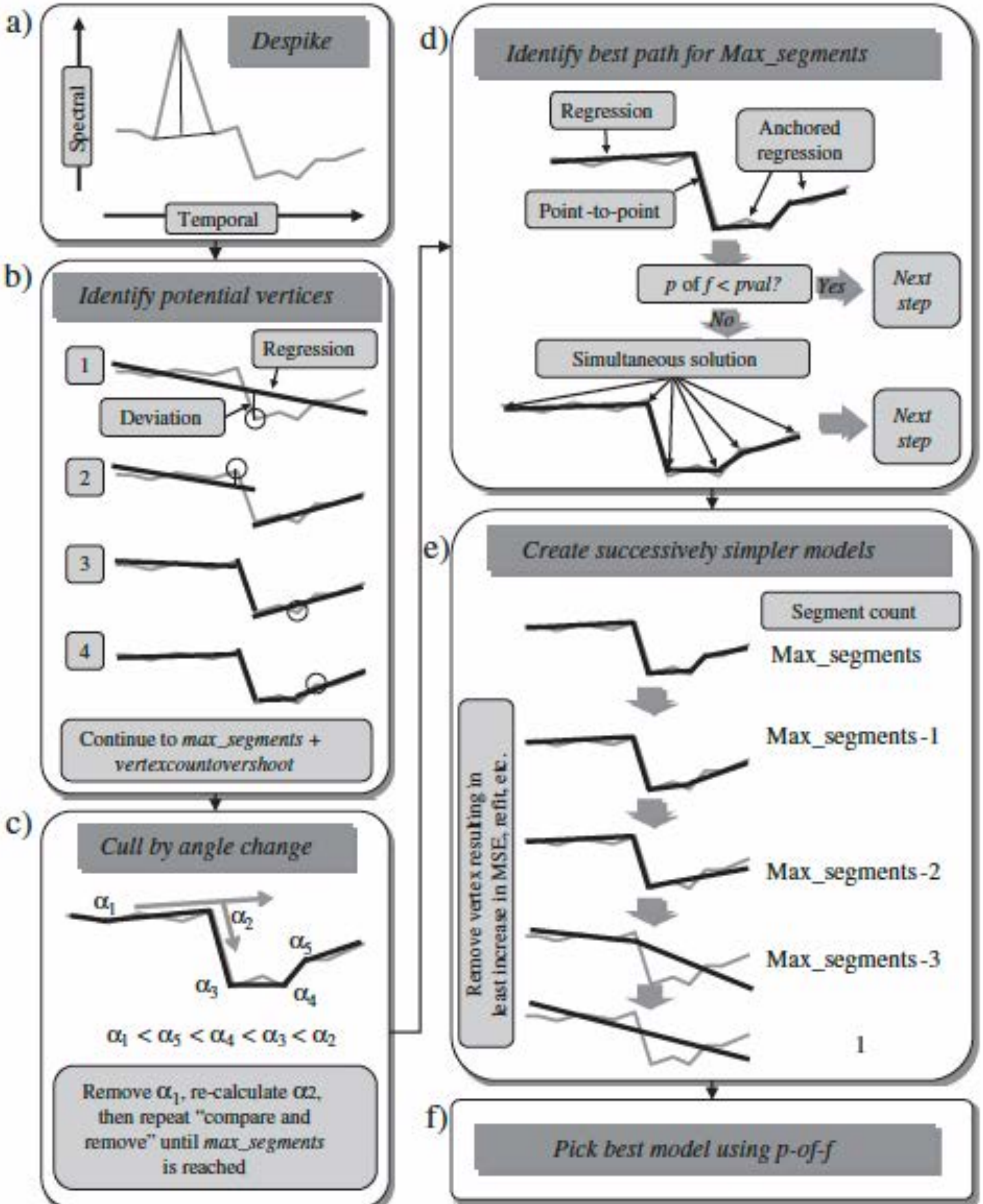
- CCDC (Continuous Change Detection and Classification)

Zhu, Z., and Woodcock, C.E. 2014. Continuous change detection and classification of land cover using all available Landsat data, Remote Sensing of Environment, 144: 152-171

LandTrendr

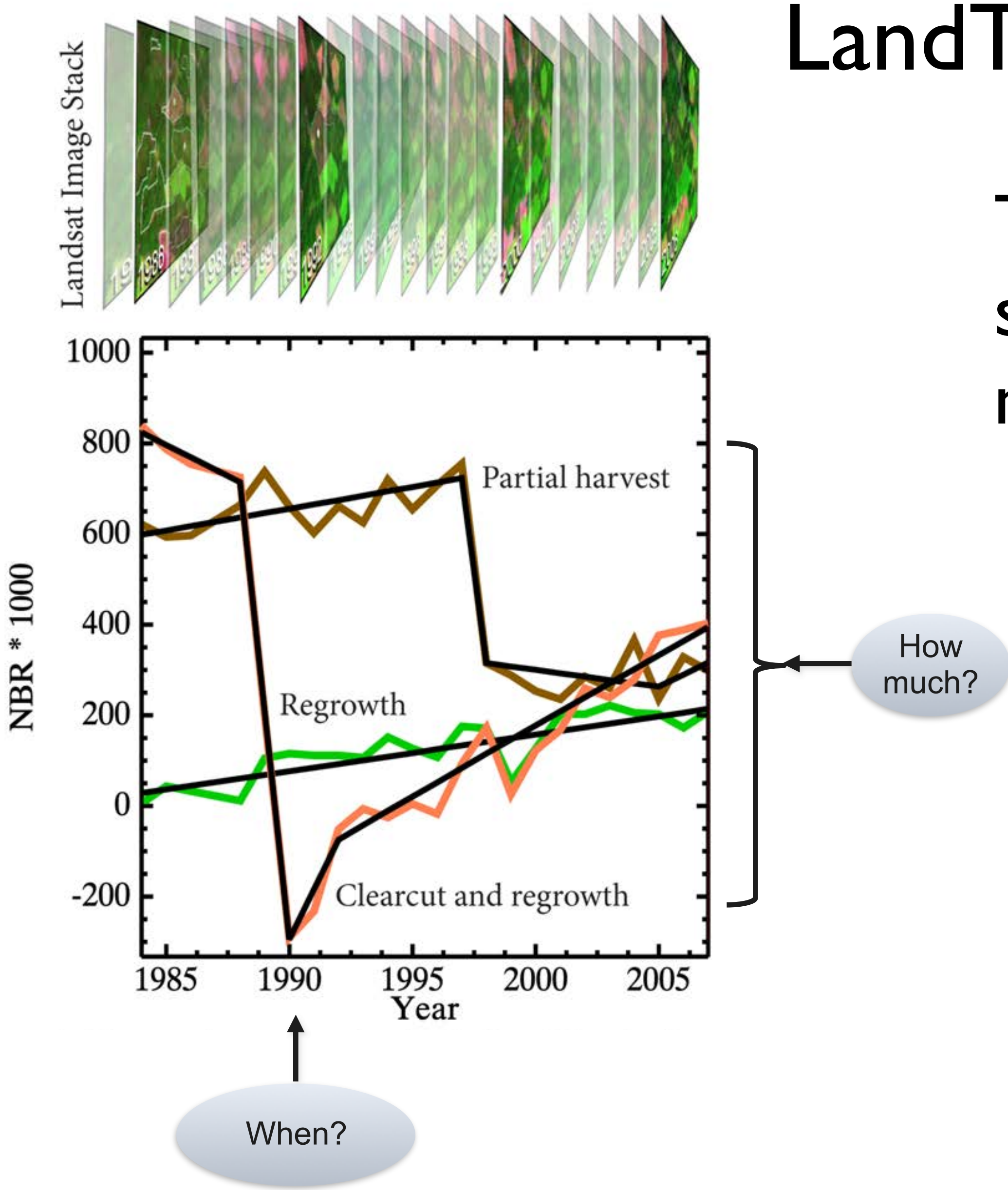
LandTrendr work flow





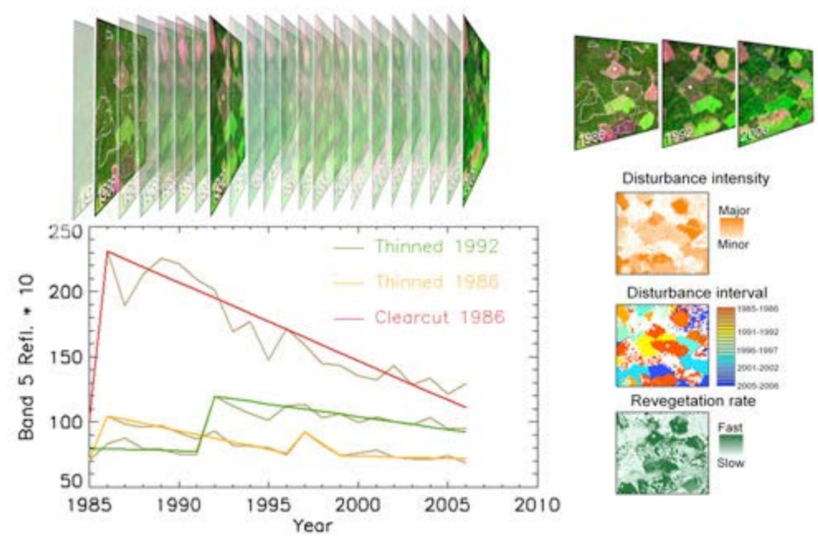
Temporal segmentation

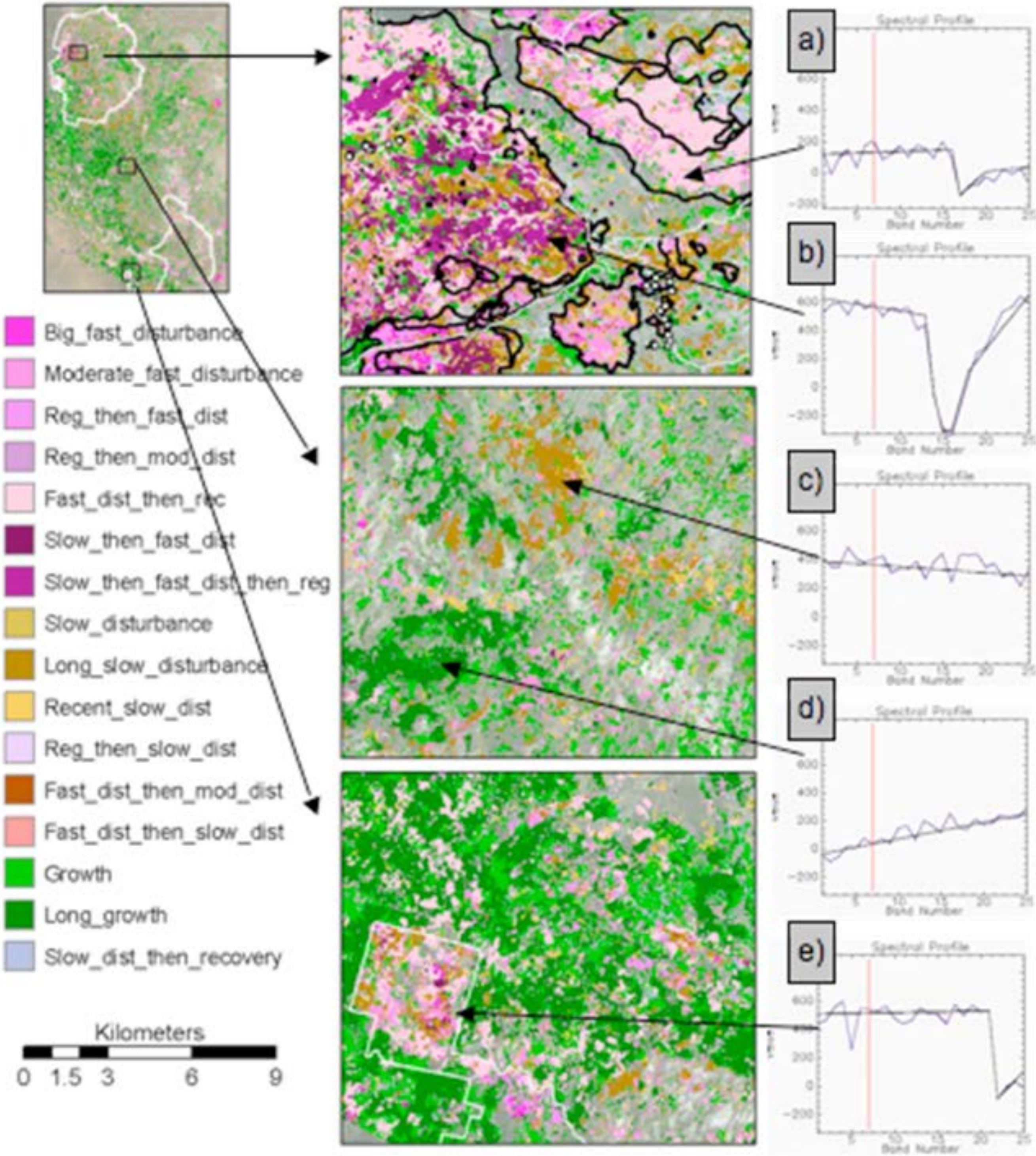
LandTrendr



Take advantage of dense time-series imagery to improve mapping

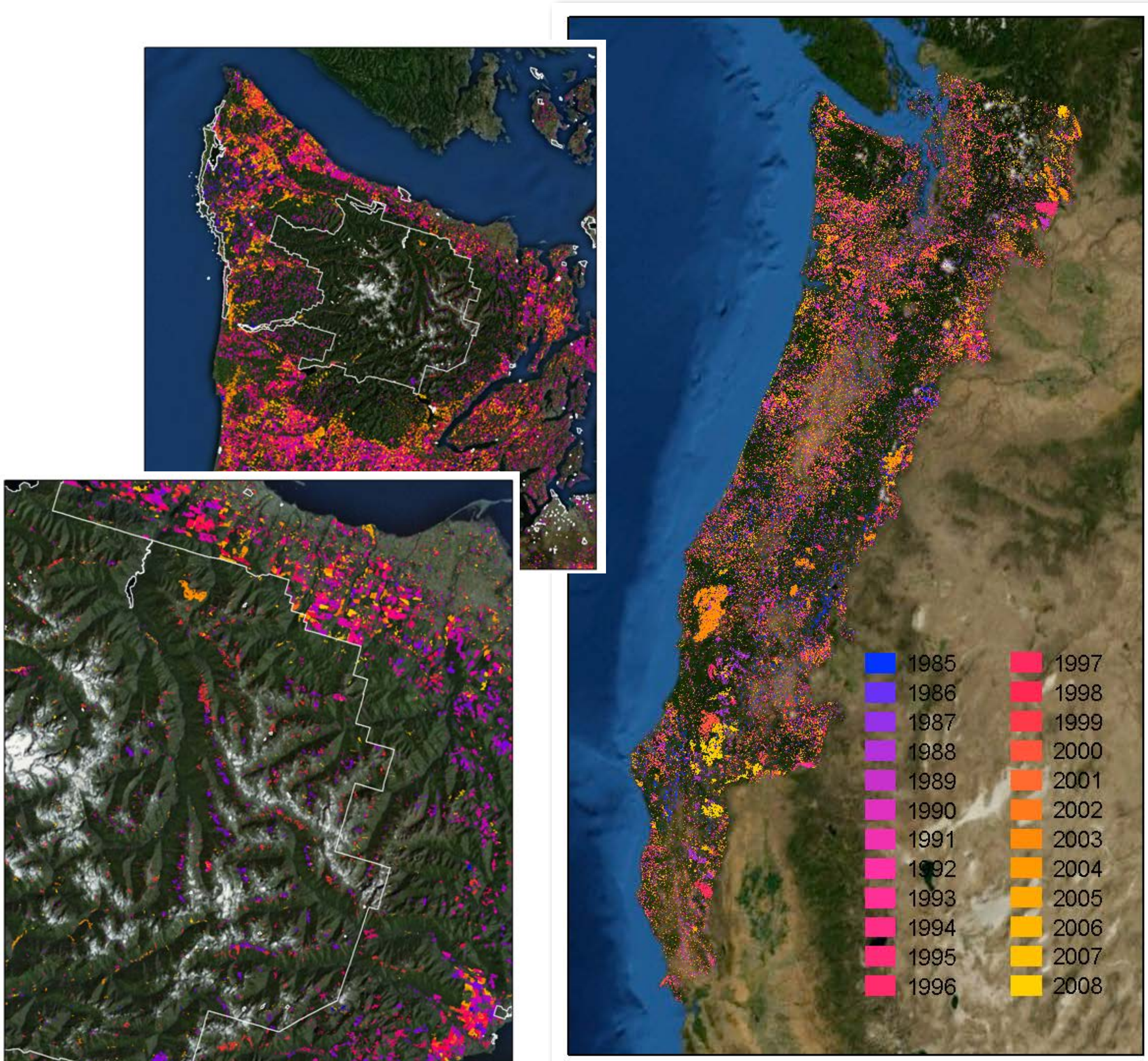
- Find both abrupt anomalies and longer-term trends
- Separate subtle changes from background noise





Mapping change

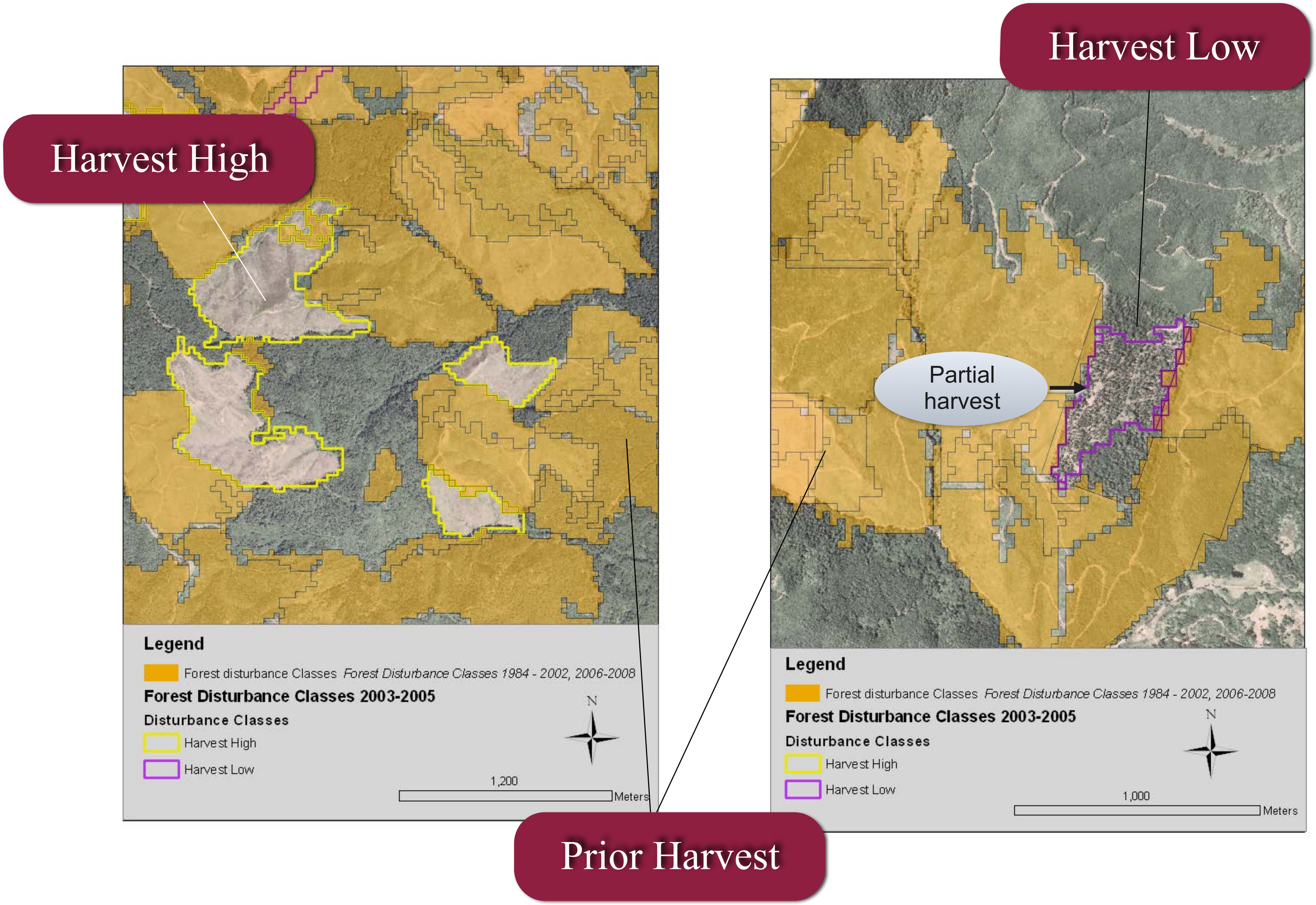
Forest Disturbance Mapping

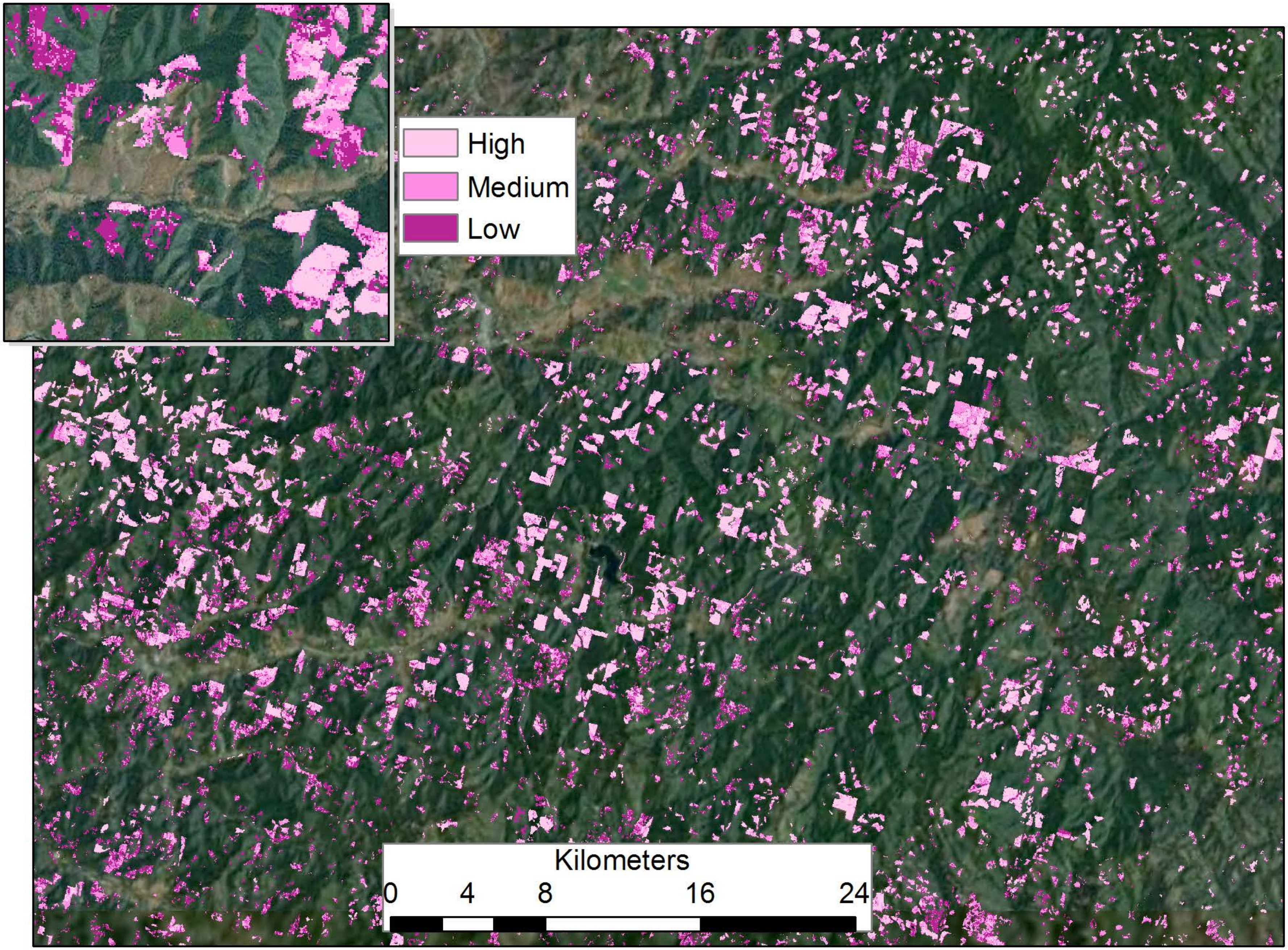


Project: Region 6 Effectiveness Monitoring Program for the Northwest Forest Plan (NWFP)

Data: > 500 individual Landsat scenes

What about intensity of harvest?



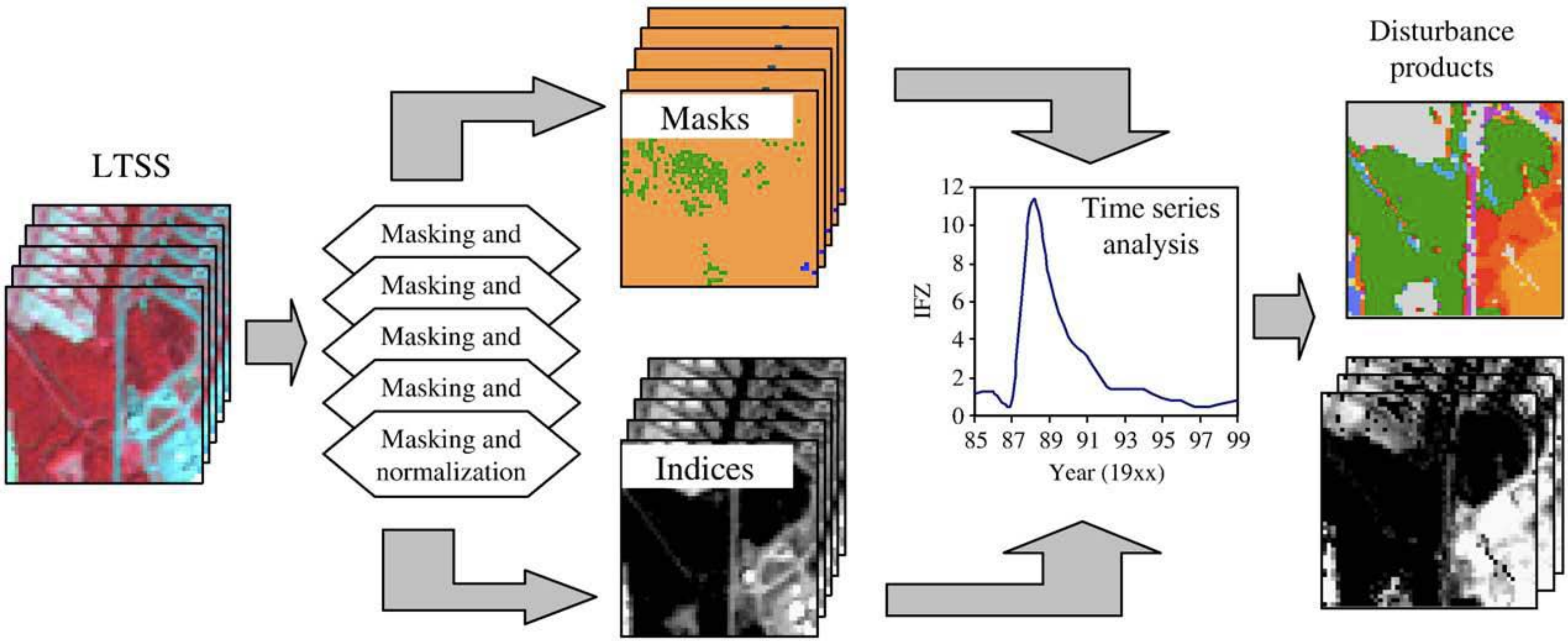

High
Medium
Low
Kilometers
0
4
8
16
24

Vegetation Change Tracker (VCT)

VCT

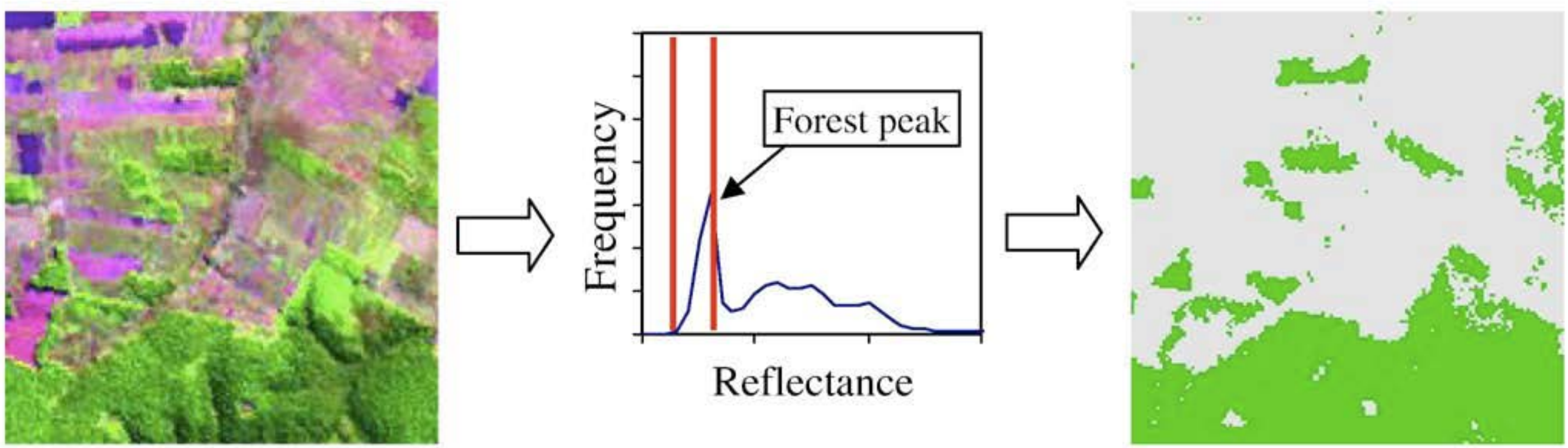
- A highly automated algorithm called vegetation change tracker (VCT) developed for reconstructing recent forest disturbance history using Landsat time series stacks (LTSS)
- This algorithm is based on the spectral–temporal properties of land cover and forest change processes, and requires little or no fine tuning or most forests with closed or near close canopy cover
- The approach has been used to examine disturbance patterns with a biennial temporal interval from 1984 to 2006 for many locations across the conterminous U.S.

How it works



LTSS
Masking and
Masking and
Masking and
Masking and
Masking and normalization
Masks
Indices
Time series analysis
IFZ
Year (19xx)
0
2
4
6
8
10
12
85
87
89
91
93
95
97
99
Disturbance products

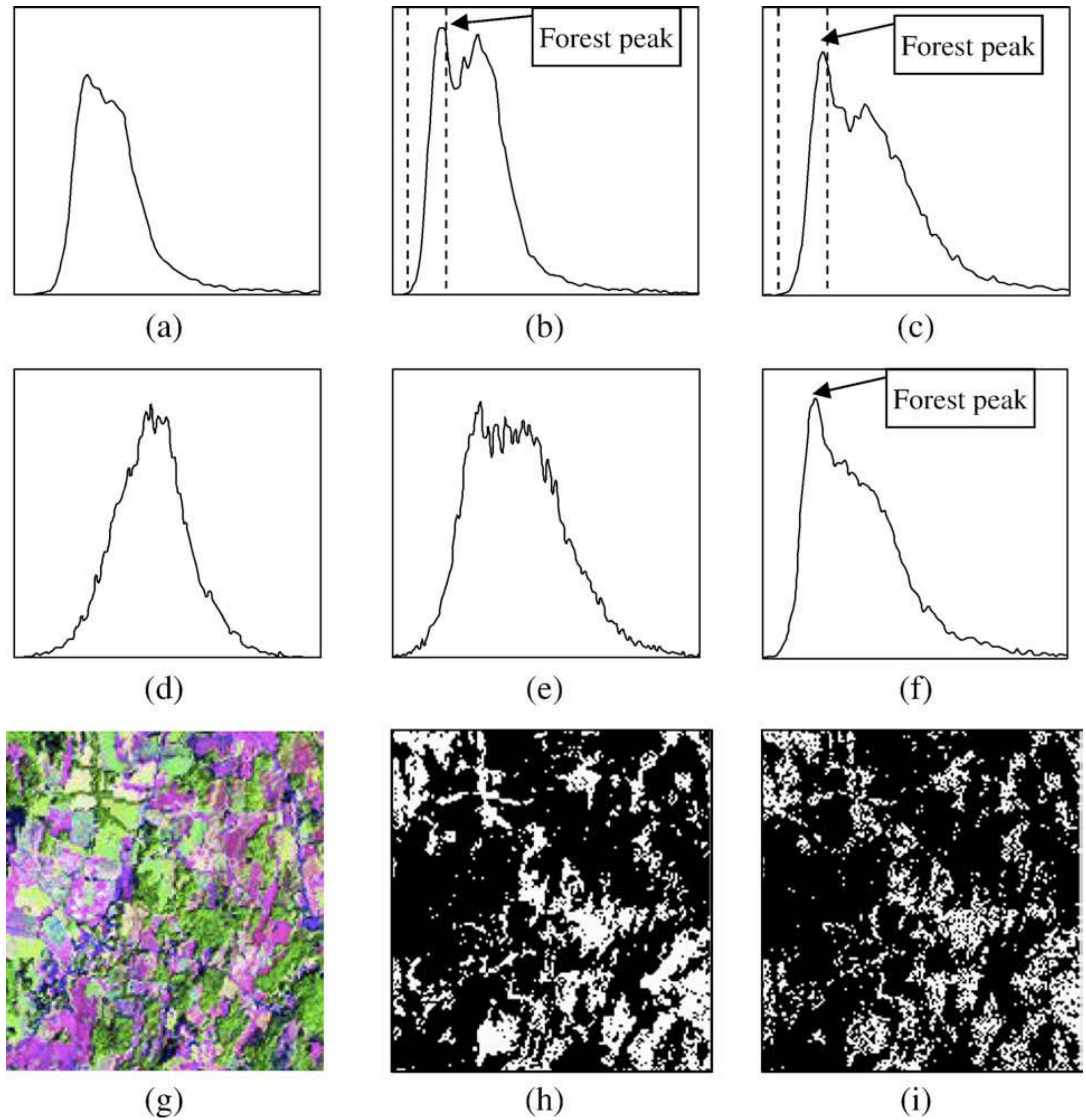
Dark forest concept

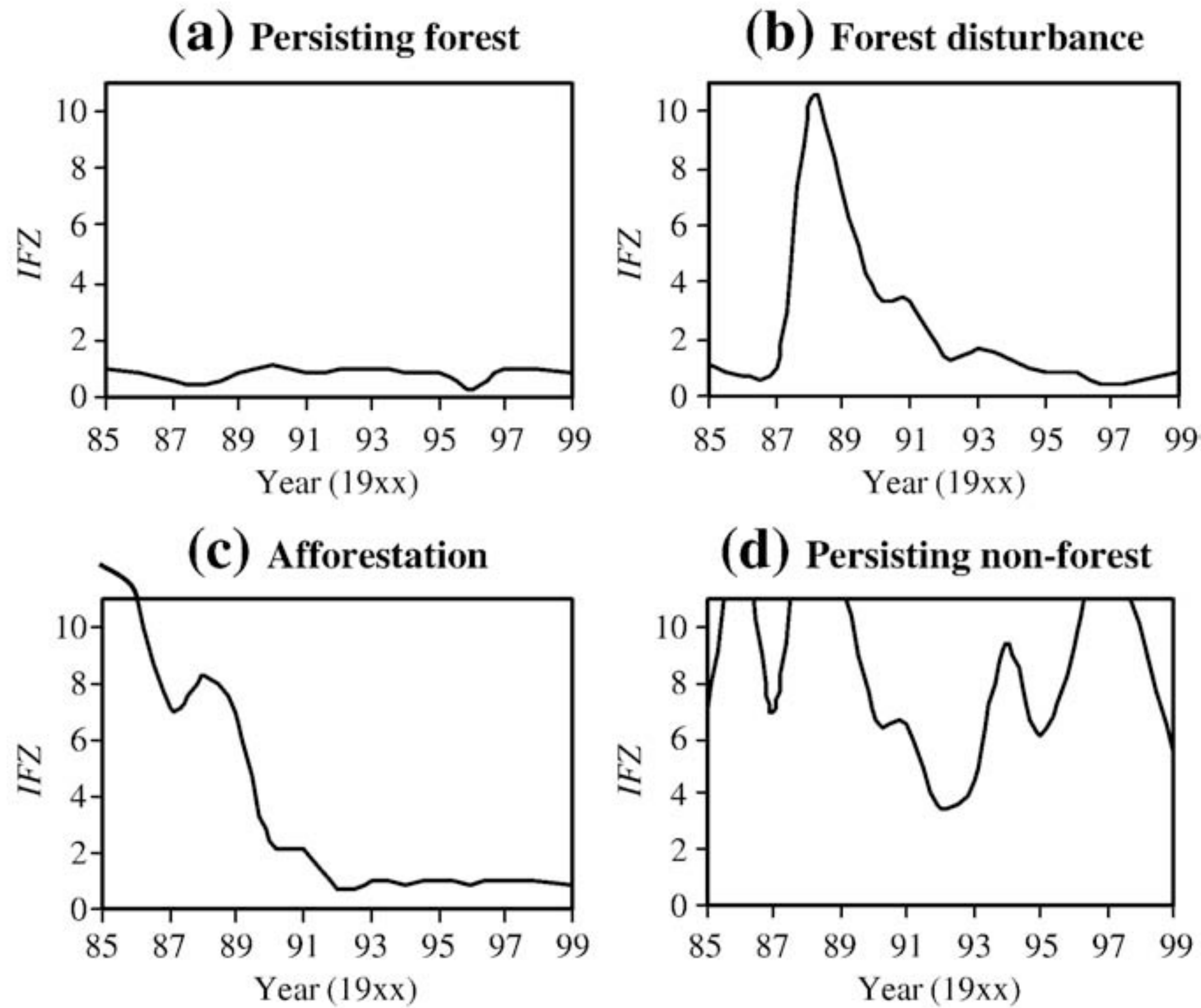


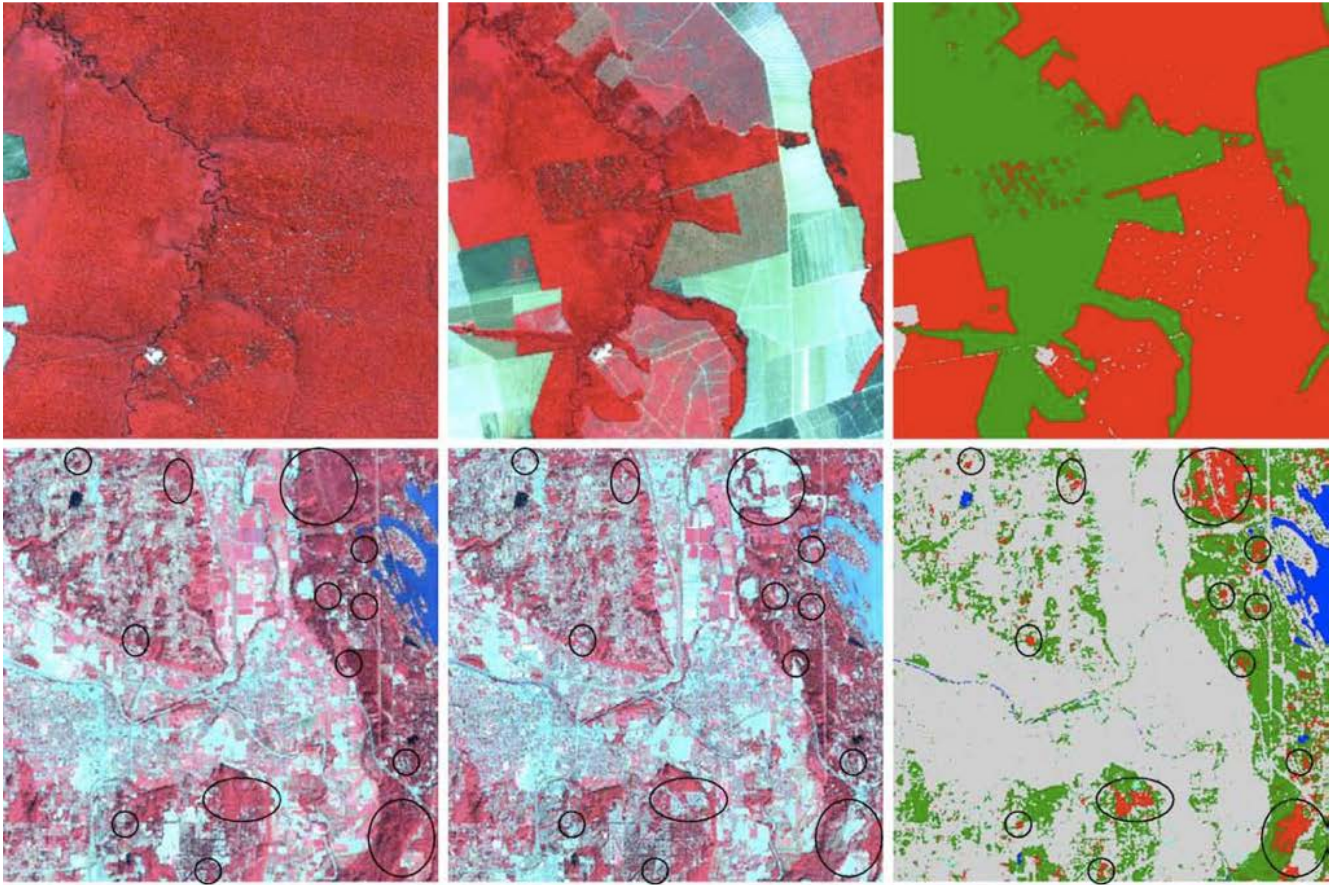
(a) Local image window (TM bands 5, 4, and 3 shown in red, green and blue)

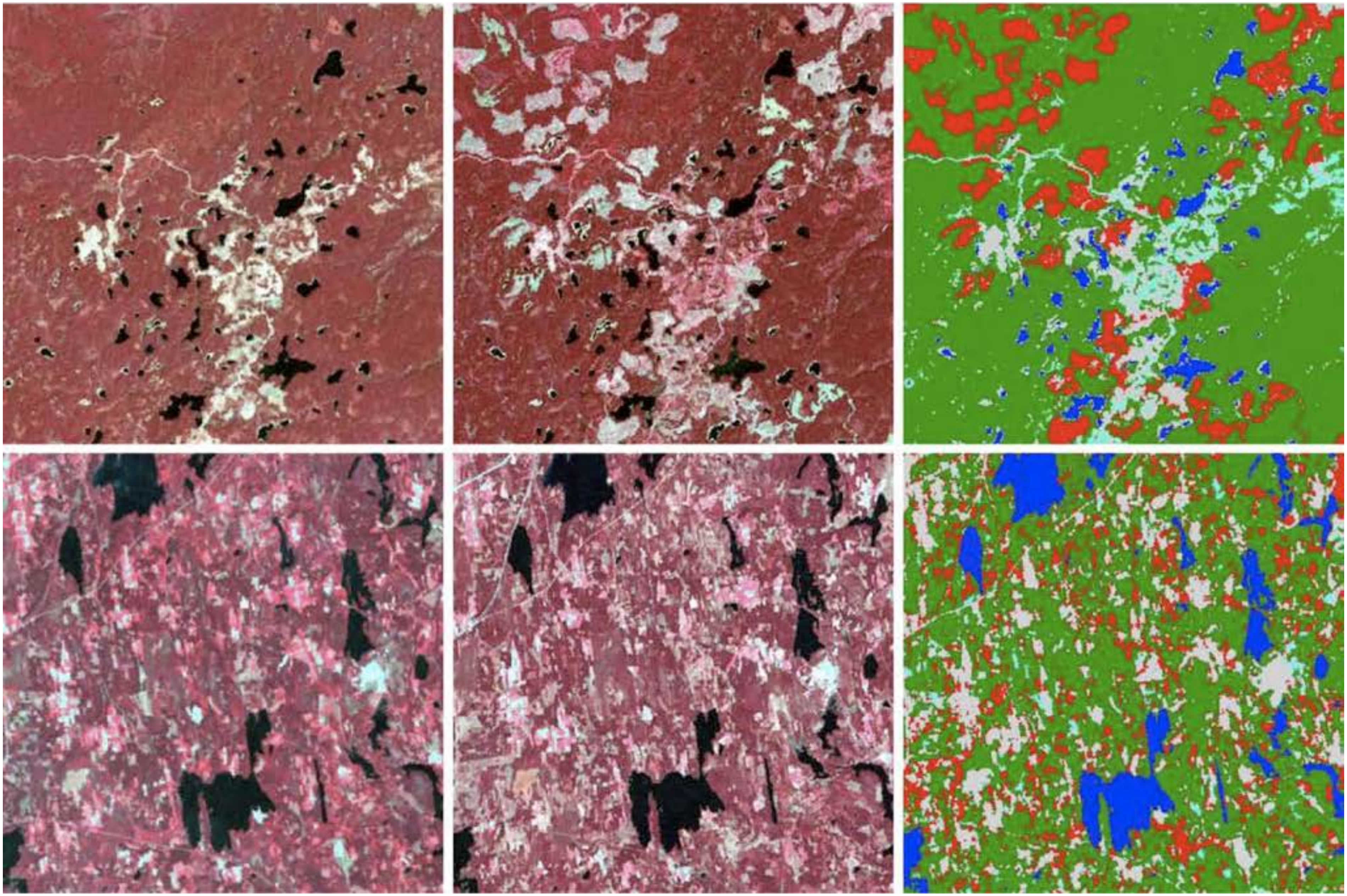
(b) Forest peak in a histogram constructed using the red band and threshold values defined by the peak (red lines)

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


Forest peak
Forest peak
Forest peak
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)


(a) Persisting forest
(b) Forest disturbance
(c) Afforestation
(d) Persisting non-forest
IFZ
0
2
4
6
8
10
85
87
89
91
93
95
97
99
Year (19xx)



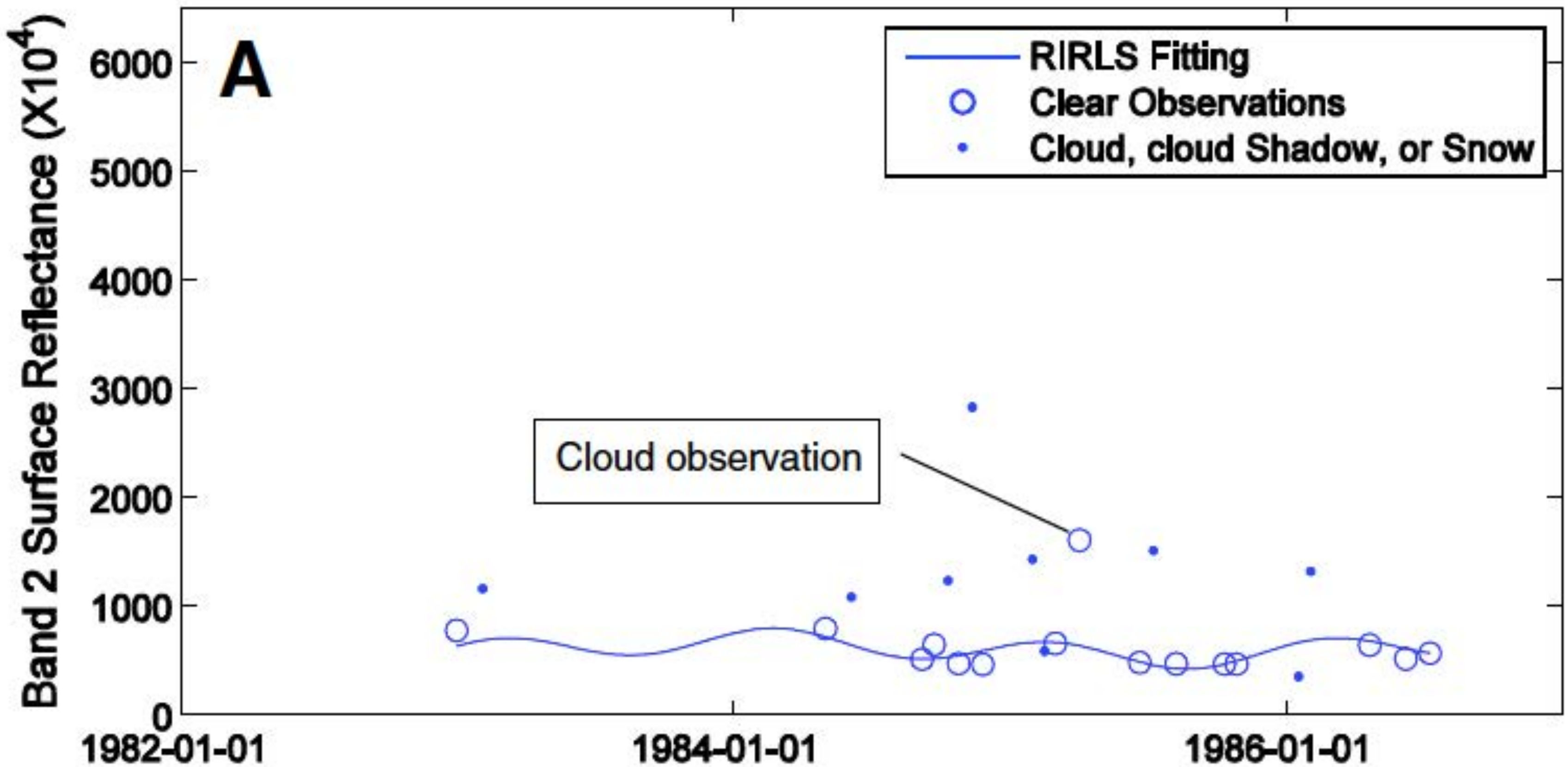


CCDC
(Continuous Change Detection and Classification)

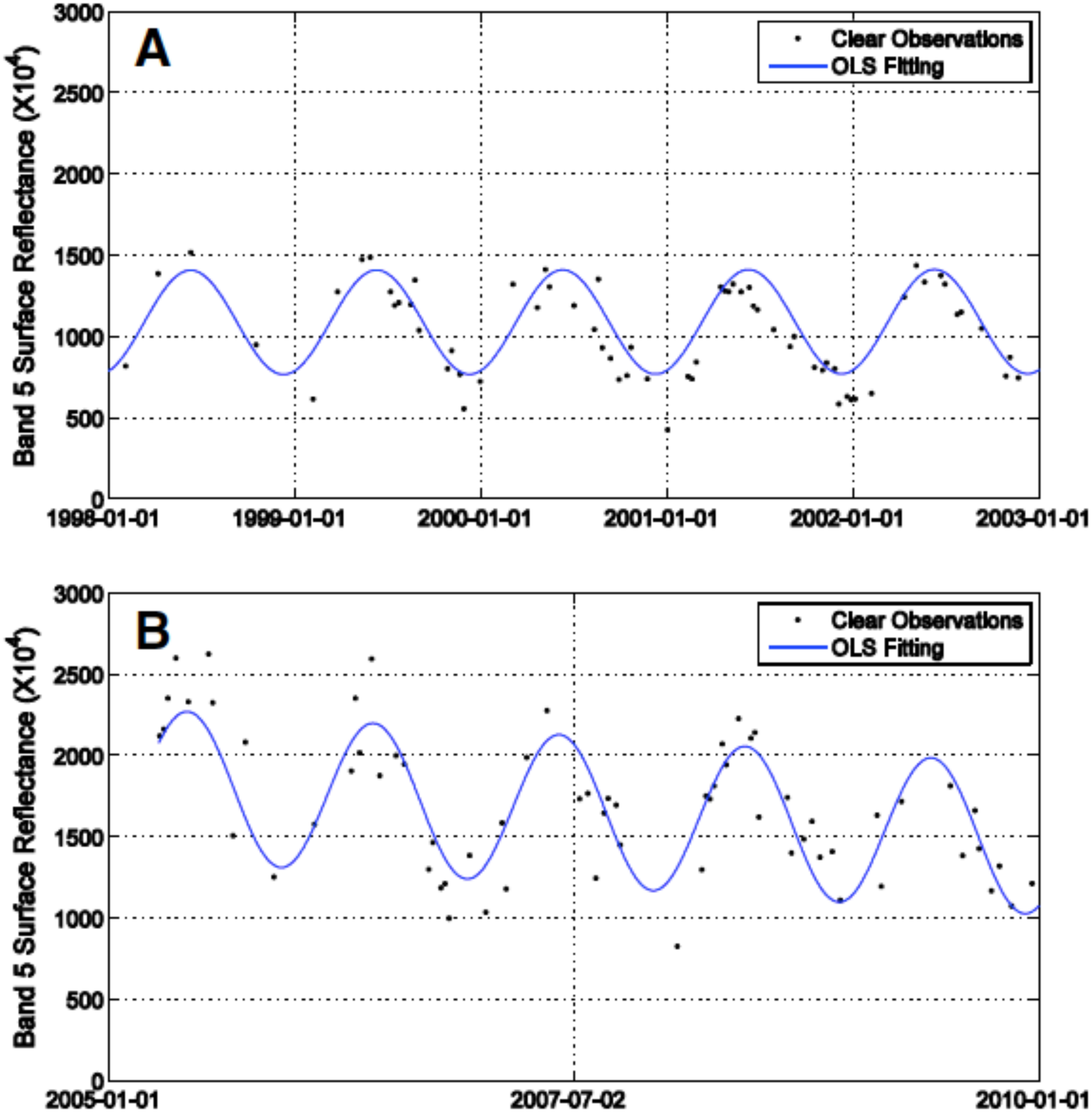
CCDC

- CCDC is based on a two-step cloud, cloud shadow, and snow masking algorithm that is used for eliminating "noisy" observations.
- Then, a time series model that has components of seasonality, trend, and break estimates surface reflectance and brightness temperature. The time series model is updated dynamically with newly acquired observations.
- Due to the differences in spectral response for various kinds of land cover change, the CCDC algorithm uses a threshold derived from all seven Landsat bands.
- When the difference between observed and predicted images exceeds a threshold three consecutive times, a pixel is identified as land surface change. Land cover classification is done after change detection

First step is to screen observations for clouds



Zhu and Woodcock (2014)



the second step is
to
fit a harmonic-
based
time series model
to
screened
observations

Zhu and Woodcock (2014)

The harmonic model

$$\hat{\rho}(i,x)_{OLS} = a_{0,i} + a_{1,i}\cos\left(\frac{2\pi}{T}x\right) + b_{1,i}\sin\left(\frac{2\pi}{T}x\right) + c_{1,i}x \qquad (3)$$

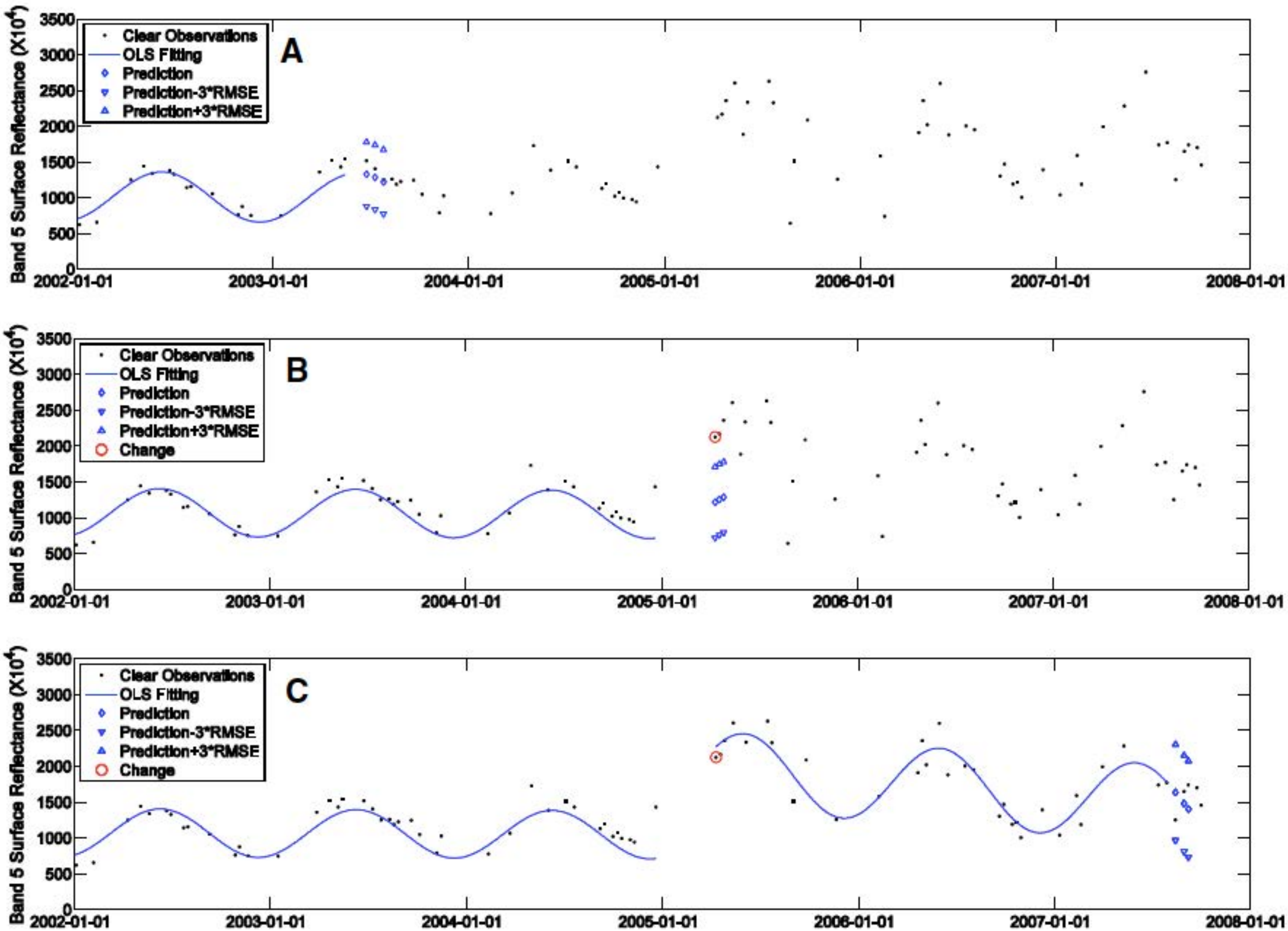
$$\{\tau^*_{k-1} < x \le \tau^*_k\}$$

where,

x	Julian date
i	the ith Landsat Band
T	number of days per year (T = 365)
$a_{0,i}$	coefficient for overall value for the ith Landsat Band
$a_{1,i}, b_{1,i}$	coefficients for intra-annual change for the ith Landsat Band
$c_{1,i}$	coefficient for inter-annual change for the ith Landsat Band
τ^*_k	the kth break points.
$\hat{\rho}(i,x)_{OLS}$	predicted value for the ith Landsat Band at Julian date x.

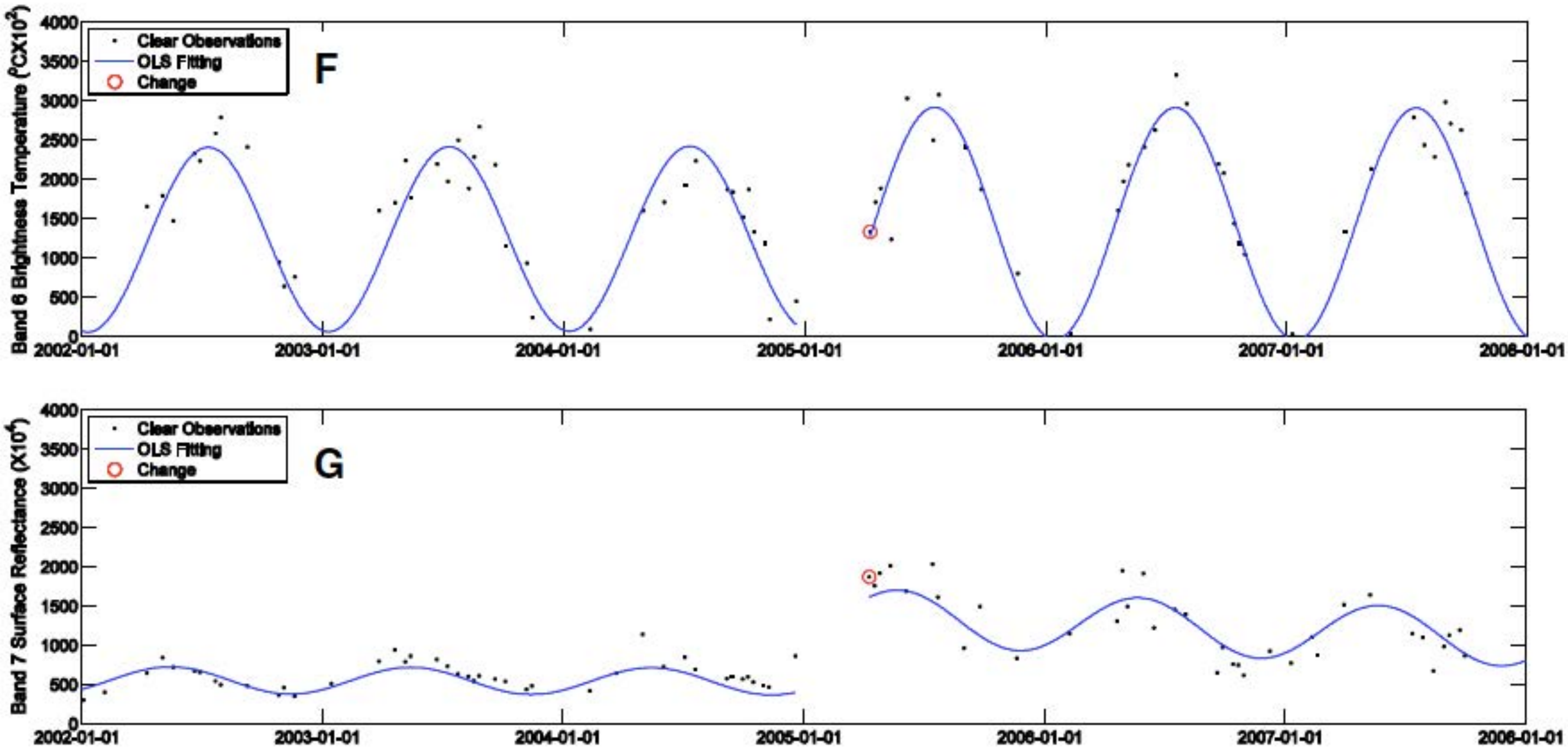
Zhu and Woodcock (2014)

Threshold-based change detection

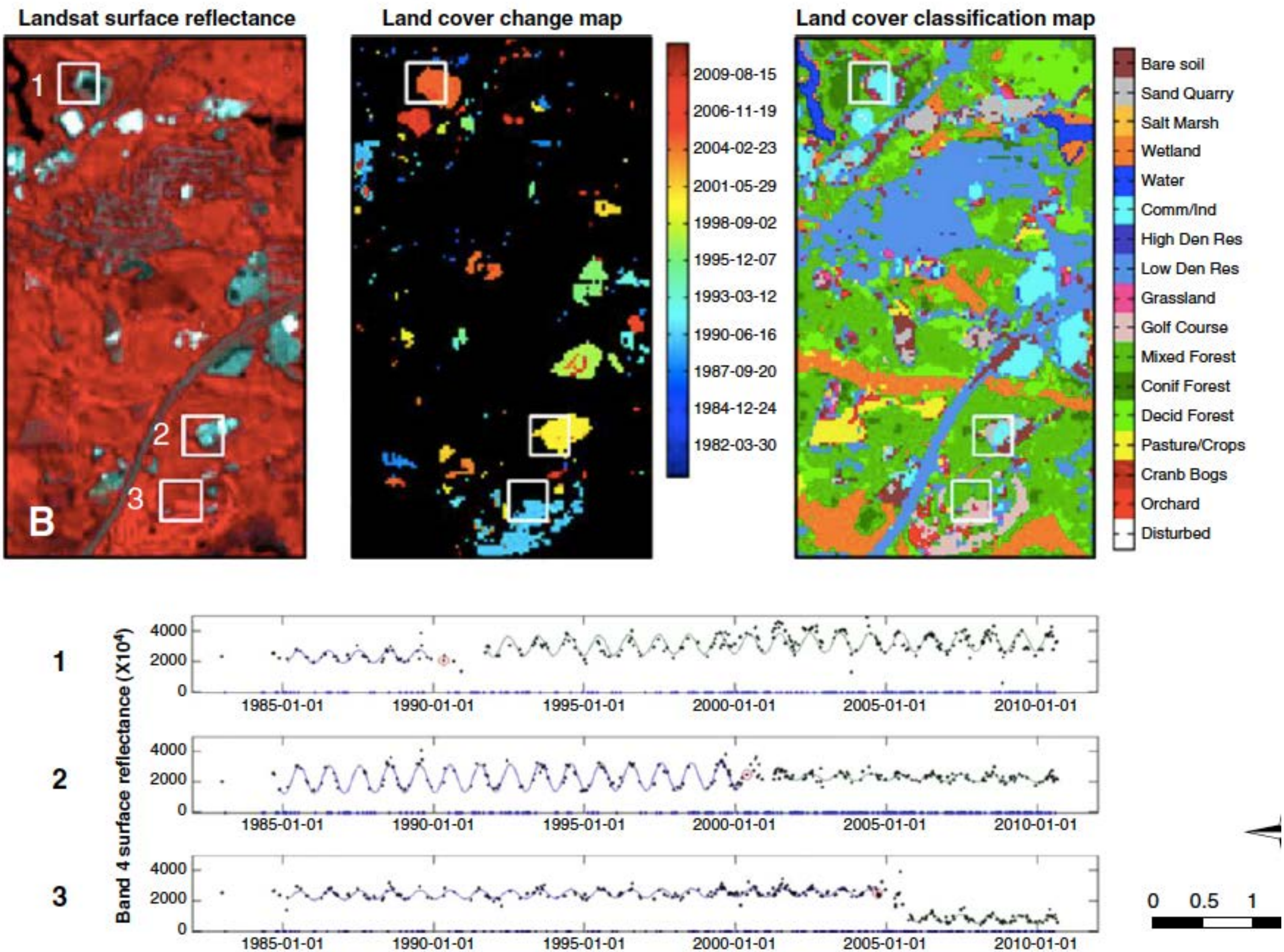


Zhu and Woodcock (2014)

Threshold-based change detection



Zhu and Woodcock (2014)



Landsat surface reflectance
Land cover change map
Land cover classification map
B
1
2
3
2009-08-15
2006-11-19
2004-02-23
2001-05-29
1998-09-02
1995-12-07
1993-03-12
1990-06-16
1987-09-20
1984-12-24
1982-03-30
Bare soil
Sand Quarry
Salt Marsh
Wetland
Water
Comm/Ind
High Den Res
Low Den Res
Grassland
Golf Course
Mixed Forest
Conif Forest
Decid Forest
Pasture/Crops
Cranb Bogs
Orchard
Disturbed
Band 4 surface reflectance (X10⁴)
4000
2000
0
1985-01-01
1990-01-01
1995-01-01
2000-01-01
2005-01-01
2010-01-01
0
0.5
1

continuous change detection

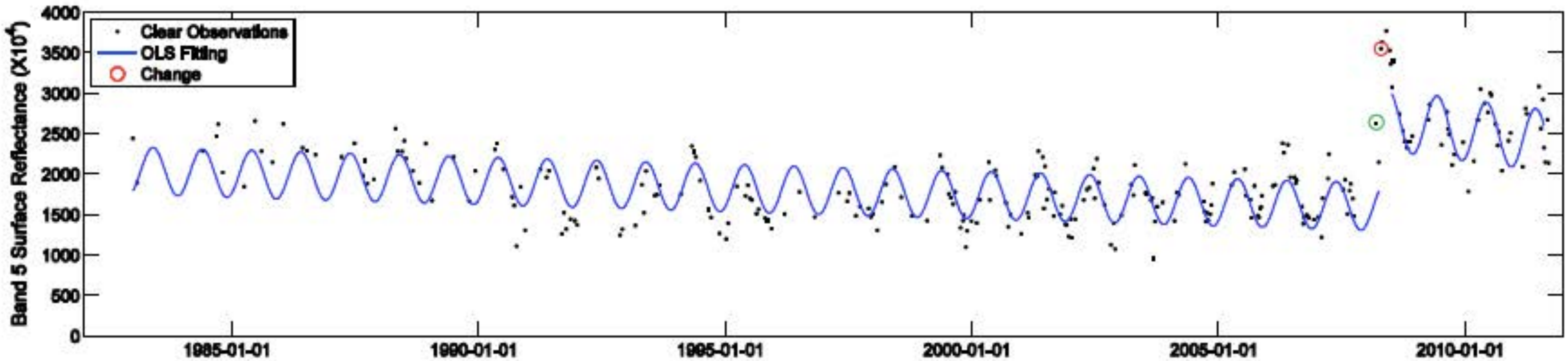


Fig. 21. Temporal error in change detection for CCDC: finding change later than it is observed in the reference data. The red circle is the "first" changed observation identified by the CCDC algorithm and the green circle is the first time the change (a partial forest cut) can be observed in close examination of the images.

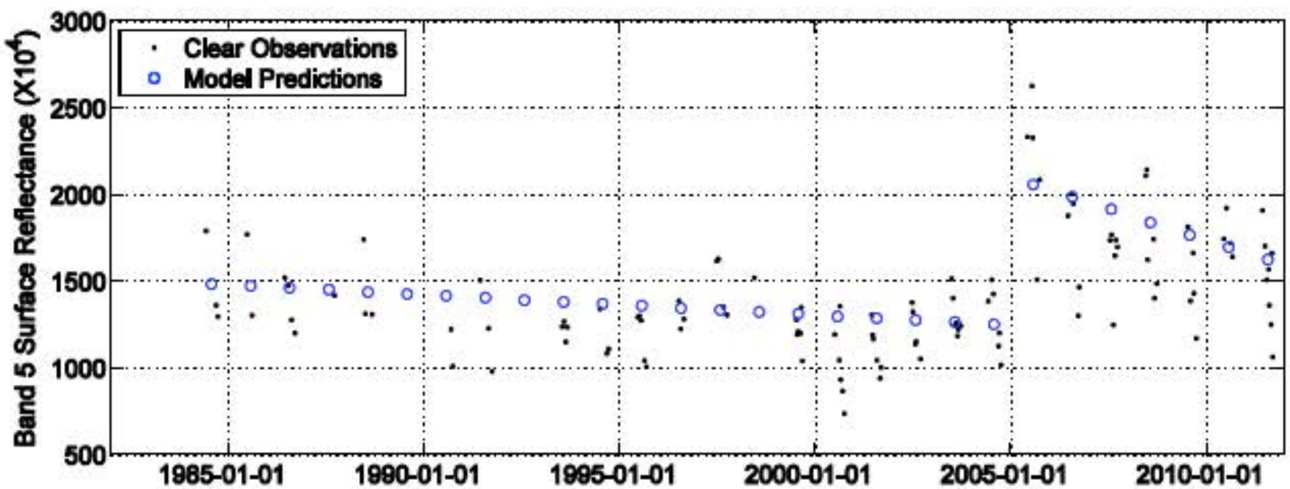


Fig. 22. Clear observations (Band 5) from Jun. to Sept. and model predictions for Aug. 1st every year between 1984 and 2011. Forest clearing occurred in 2005 for this pixel.

Zhu and Woodcock (2014)

Summary

- Time series analysis of remotely sensed data use the temporal dimension of observations
- There are many established methods - mostly borrowed from chemistry, physics, and financial applications
- In remote sensing applications, we treat each pixel as individual observation series and process millions of pixels the same way (high computation requirement)
- In remote sensing applications, often prediction is not done

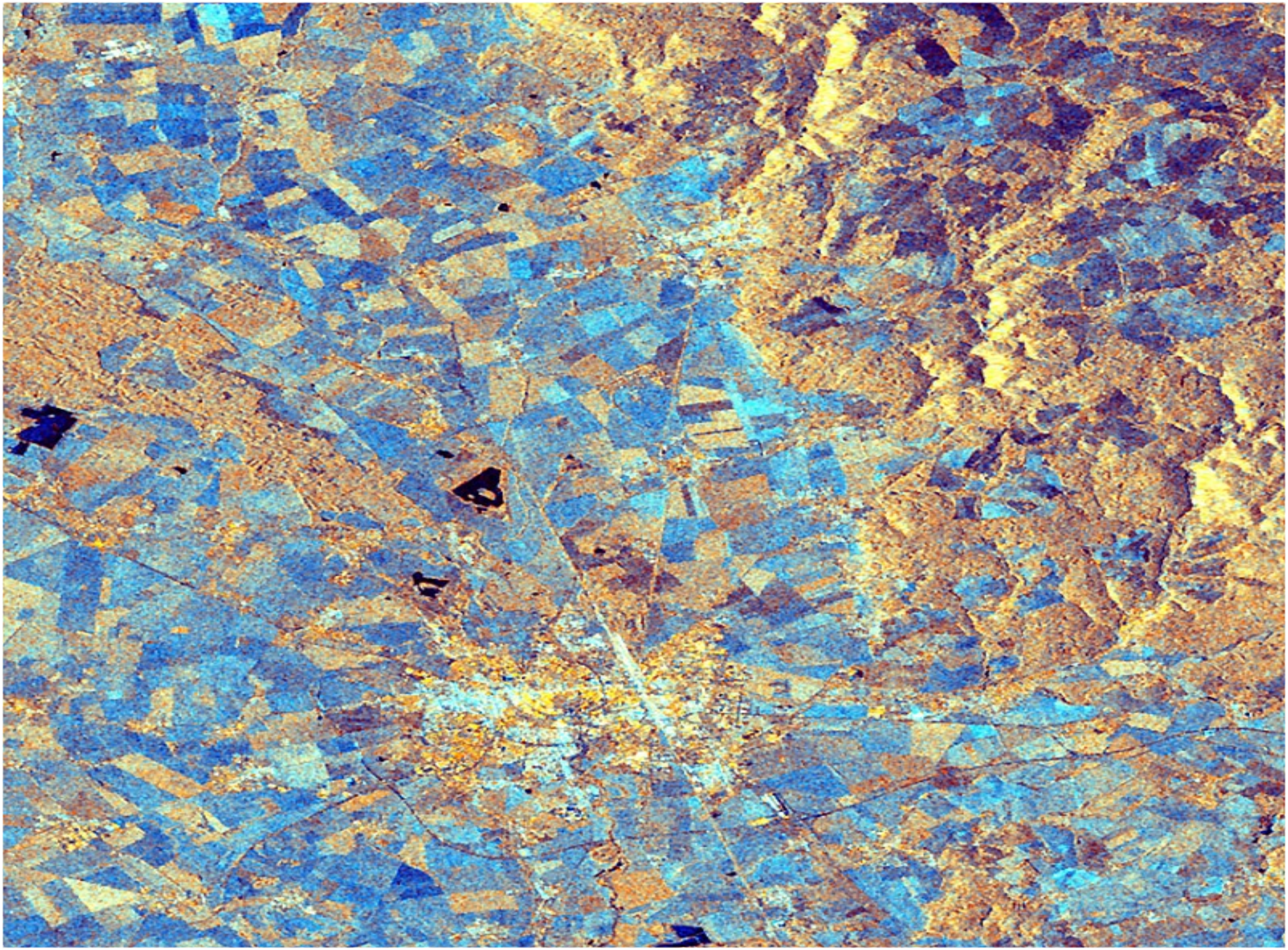
Summary

- Time series methods are often divided among the types of remotely sensed observation available
- High temporal frequency observations (e.g. MODIS, AVHRR) are easily processed with traditional time-series analysis methods
- Low temporal frequency observations (e.g. Landsat, Sentinel) significant amount of pre-processing is required and the traditional time series methods are not always applicable

Useful Links

- http://www.wageningenur.nl/en/Expertise-Services/Chair-groups/Environmental-Sciences/Laboratory-of-Geo-information-Science-and-Remote-Sensing/Research/Integrated-land-monitoring/Change_detection_and_monitoring.htm

- http://bfast.r-forge.r-project.org



Accuracy assessment

Image classification process

define the problem

select class labels

acquire data

process data to extract thematic information

perform accuracy assessment

distribute results

What is accuracy assessment?

- When a map (LC or LCC) is derived from remotely sensed data, that map is considered to be only a hypothesis!
- As with other hypothesis-based problems, the hypothesis has to be tested with data
- Testing is done by extracting samples from the map, compare these samples to a known reference
- Then accuracy can be reported using a variety of metrics and a degree of confidence can be attached to the results

What is accuracy assessment?

- When a map (LC or LCC) is derived from remotely sensed data, that map is considered to be only a hypothesis!
- As with other hypothesis-based problems, the hypothesis has to be tested with data
- Testing is done by extracting samples from the map, compare these samples to a known reference
- Then accuracy can be reported using a variety of metrics and a degree of confidence can be attached to the results

Common questions in accuracy assessment

- What is my sample and how do I extract it?

- A sample is a subset of a population - in general, the population (the pixels in a map) is very large, making a census of all the values impractical or impossible

- The sample represents a subset of manageable size

- Samples are collected and statistics are calculated from the samples so that one can make inferences or extrapolations from the sample to the population (the map) with known confidence intervals

- This process of collecting information from a sample is referred to as sampling

How do you collect reference data?

- Back-classification of training data
  – Use your own training data to see how well you did!

- Cross-validation
  –Don't make a map, use training data splits

- Independent (spatially) non-random samples
  –systematic, clustered, multi-stage etc.

- Independent (spatially) random samples
  – either stratified or not stratified (based on class value)

What about the sample size?

Sample size depends on a number of factors including:

- Expected accuracy
- Desired accuracy
- Desired level of confidence interval
- <u>ultimately the resources available (accuracy assessment is expensive!)</u>

A common approach is to decide on the total sample number first and allocate based on the problem

Sample size conundrum

As little as possible   but …

as many as necessary

How to allocate?

- Proportional allocation: allocate samples based on the size (area) of the map category

- Proportional allocation works well in situations where the categories have sizable representation on the map

- It does not work well in change detection studies as the most important class (the change category) often the smallest (area-wise) category on the map

Common questions in accuracy assessment

- What is my sample and how do I extract it?

- What is my metric to measure accuracy?

Tools and metrics to use

- For categorical outcomes (i.e. classification or change detection) confusion matrix is the standard tool
- Report all accuracies (overall, class-specific, omission/commission, kappa etc.)
- For continuous outcomes (i.e. forest fraction) various statistical tools (correlation, goodness of fit etc.) can be used
- <u>The most important thing to remember is to be transparent and show all of the work and the data!</u>

Common questions in accuracy assessment

- What is my sample and how do I extract it?
- What is my metric to measure accuracy?
- Do I need to worry about spatial autocorrelation?

Spatial auto-correlation

- Spatial auto-correlation is correlation that occurs when near things are more related than far things
- Can negatively impact statistical analyses (accuracy assessment)
- The most important effect is that it inflates accuracy results!!
- Estimate correlation length (variogram) and choose samples beyond this correlation length
- In general, only applied in situations with a large sample - if have small sample - can't be picky!

Common questions in accuracy assessment

- What is my sample and how do I extract it?
- What is my metric to measure accuracy?
- Do I need to worry about spatial autocorrelation?
- What is a good accuracy?

What is a good accuracy?

- Well that depends!
- Obviously we all want the products with the highest/best accuracy
- Inverse relationship between the categorical detail and accuracy (more categories - expect reduced accuracy)
- Class specific accuracies matter as much as the overall accuracy of the product
- Change detection problems are inherently more difficult so perhaps lower expectation

Things never to do

Things <u>not</u> to do

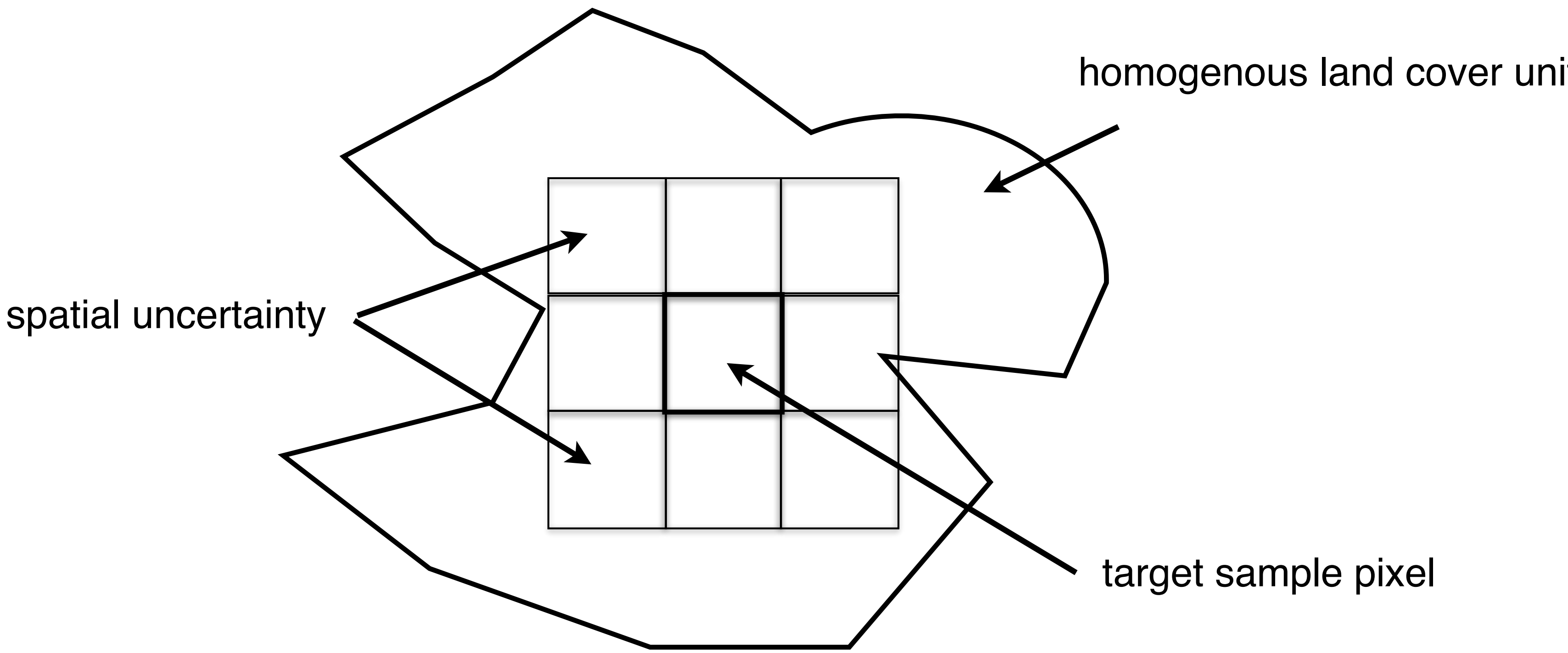
- Not assess accuracy
- Use training data in accuracy assessment (maybe used as an initial check)
- Use products without known accuracy
- Deliver products without known accuracy
- Use unreliable reference data
- Use single pixels (especially at higher spatial resolution) to perform accuracy
- Use small sample sizes - especially for important (but small) categories

Things <u>we need</u> to do

- Always perform accuracy assessment
- Report/demand validation results
- Report all accuracies and show/share sample data (transparency is important)
- Use accuracy assessment results to report corrected area estimates
- Use area-corrected accuracy assessment tools
- Allocate more resources to accuracy assessment

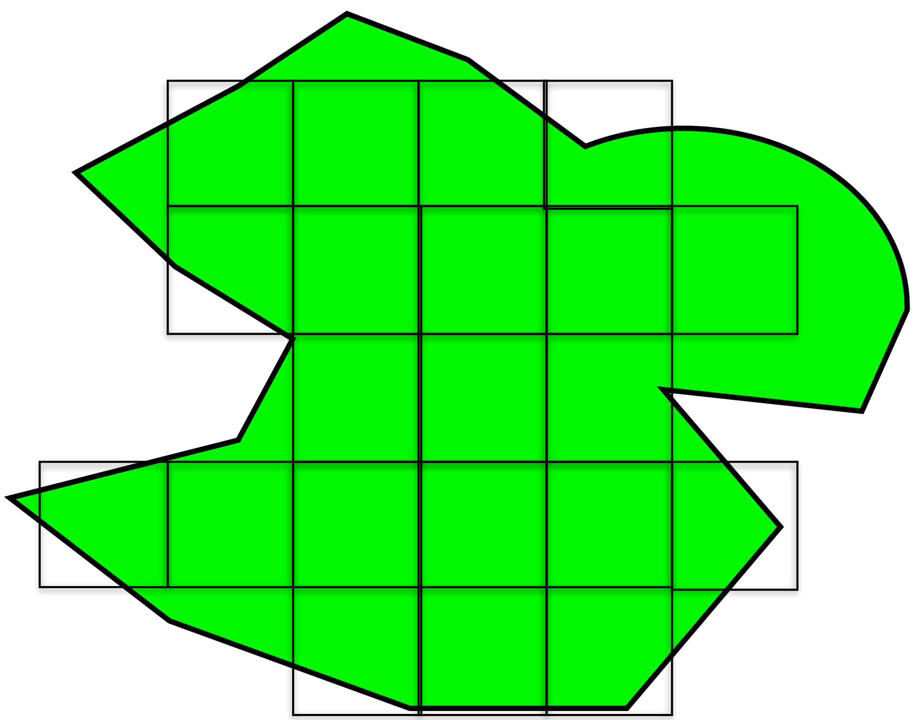
Loose ends in accuracy assessment (1)

- Single pixel accuracy assessment is problematic - especially in higher spatial resolution products
- Due to unknown location of the pixel (image registration always has problems)



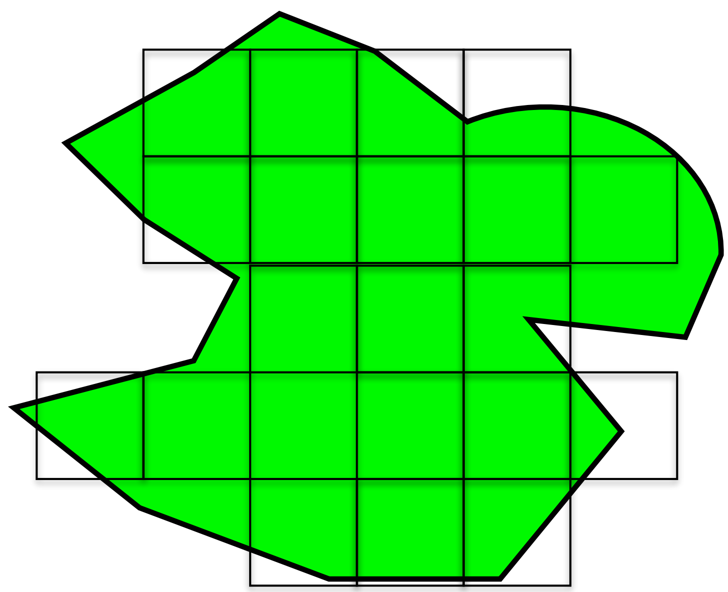
Loose ends in accuracy assessment (2)

- In site (region or polygon) based accuracy - we must sample the polygon and not the pixels in polygon
- Must find a way to interpret the polygon in the reference data (majority rule?)



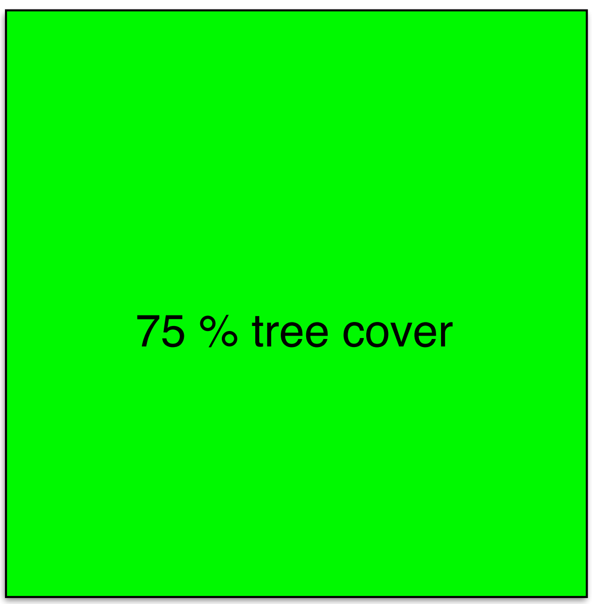
Loose ends in accuracy assessment (3)

- Must separate pixels from regions in cross-validation
- If you are using regions (polygons) as the map unit, you cannot use individual pixels from a region to do the cross-validation based accuracy assessment - regions and pixels must be kept together as either training or testing groups



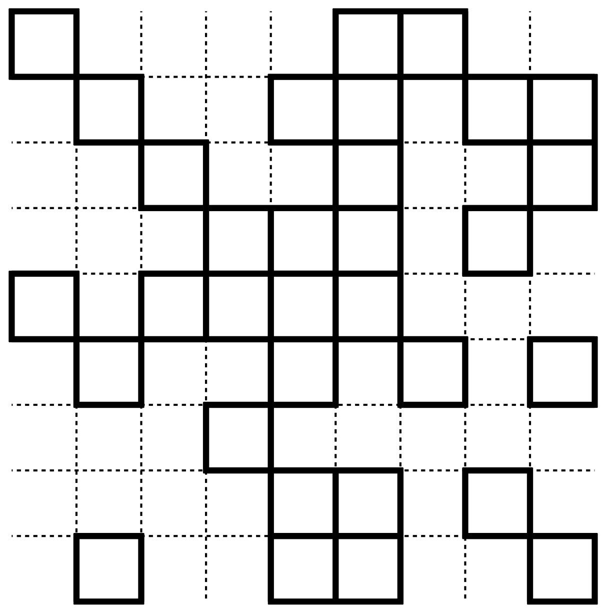
Loose ends in accuracy assessment (4)

- For continuous outcomes (i.e. tree fraction) we need to generate continuous reference sets

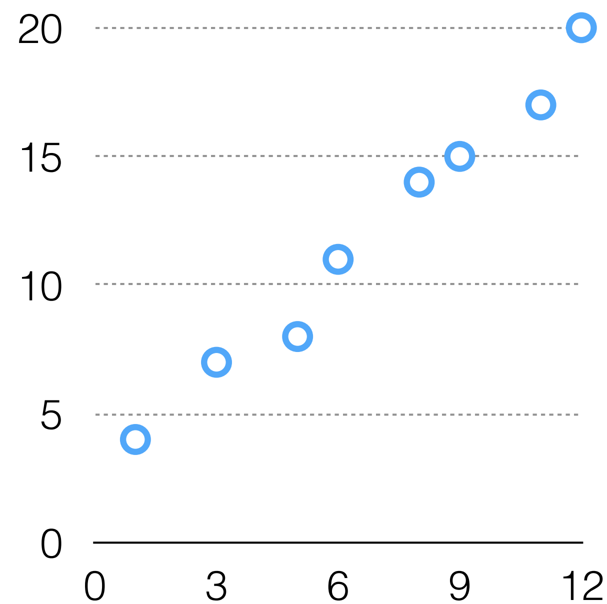


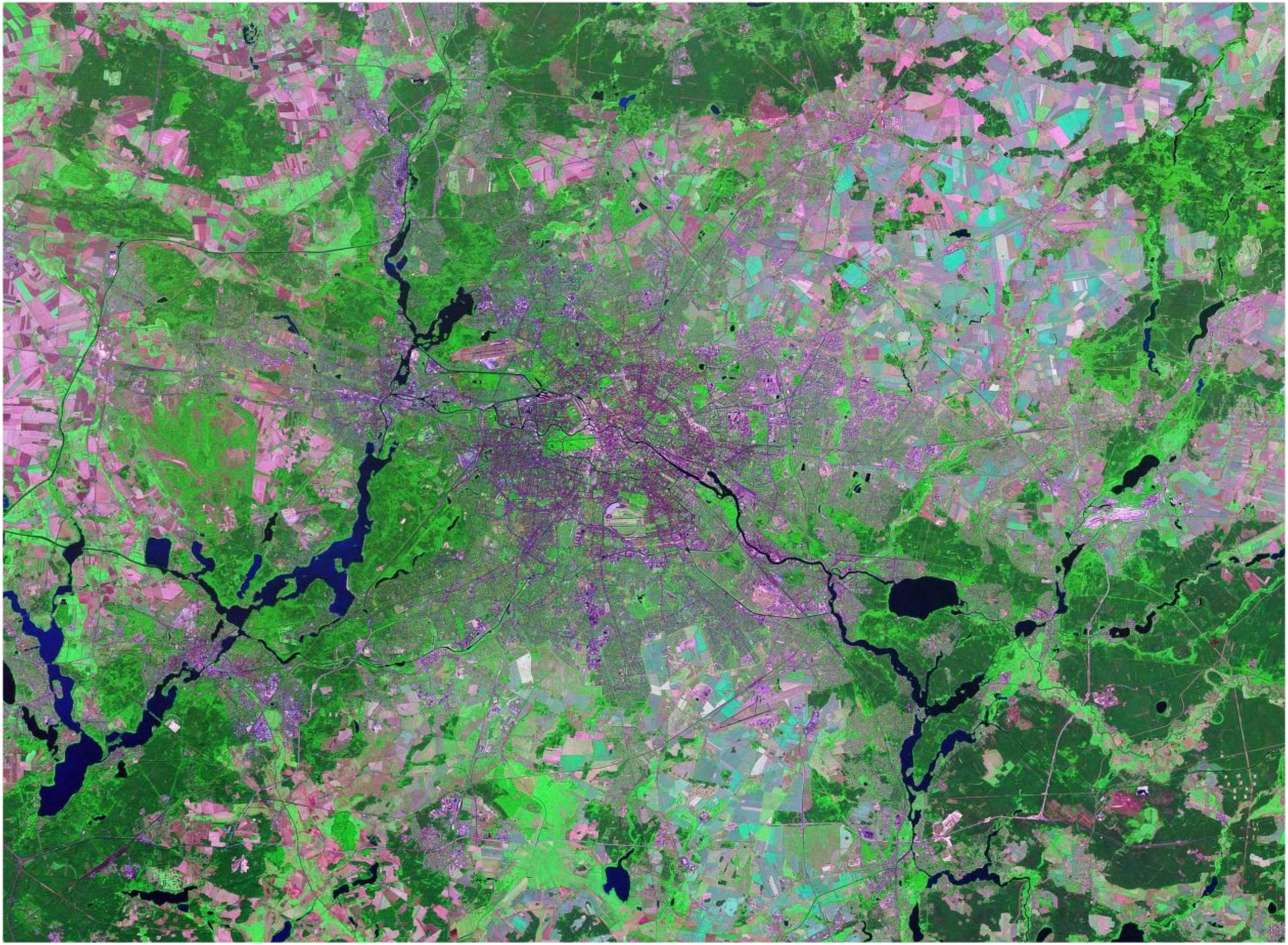
30 meter pixel

reference high resolution map



32/81 pixels (~ 40% tree cover)





Summary

- Accuracy assessment is important but expensive
- Must incorporate accuracy assessment into original planning - can't be treated as an afterthought
- Need to report all the work and (if possible) make sample data available (transparency is key)
- Need to move away from single pixel assessments (especially for high resolution products)
- There are tools to correct area estimates based on accuracy assessment as well as tools to perform area adjusted accuracies
- As a community we need to make accuracy assessment a priority

Thank you

ozdogan@wisc.edu