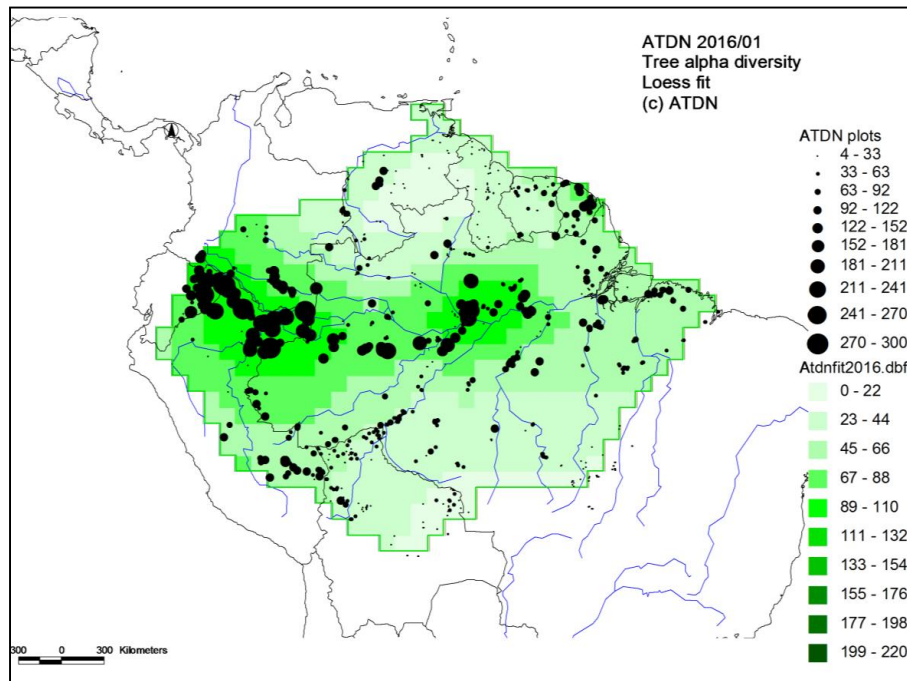# Mining the Landsat archive for detecting spatial and temporal vegetation patterns in Amazonia

Jasper Van doninck

University of Turku
Amazon Research Team
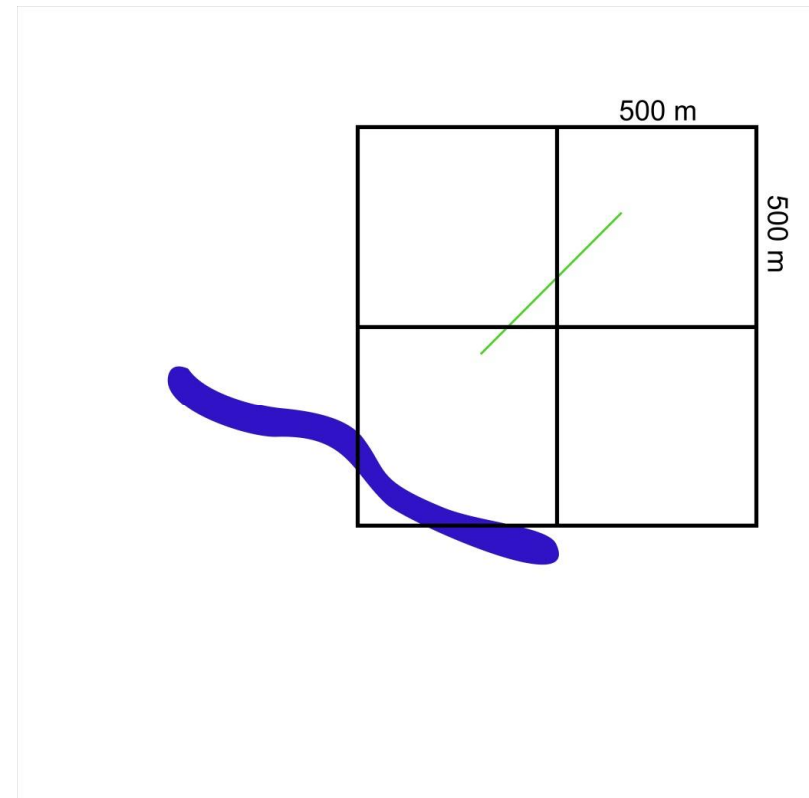Department of Geography and Geology / Department of Biology

UNIVERSITY OF TURKU

Amazonian biodiversity

Low resolution (500 m-1 km) satellite data

+ Near-daily global coverage

+ Continental to global scale studies

+ Studies on vegetation phenology

- Spatial resolution mismatch with most field inventories
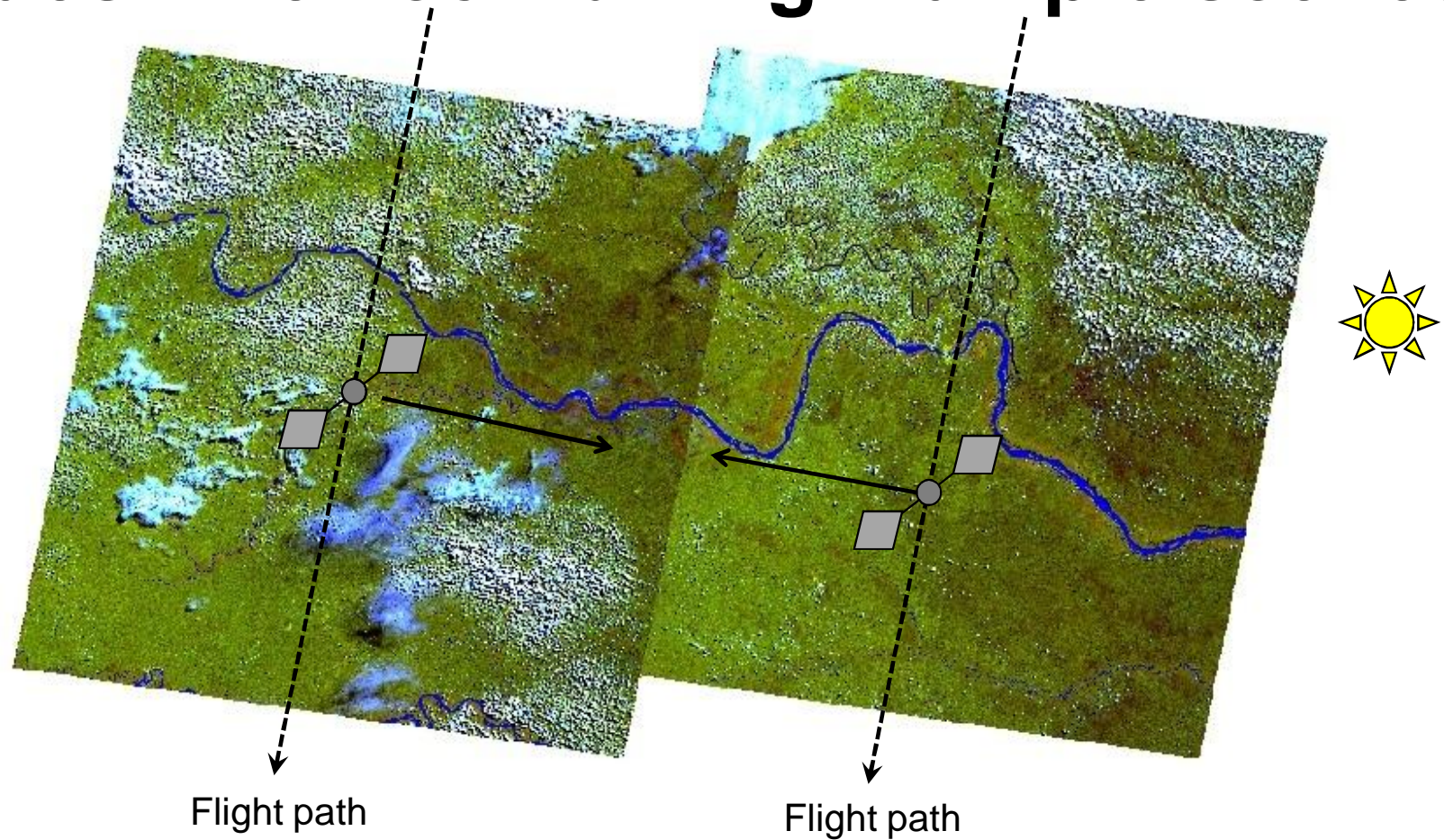
500 m

500 m

UNIVERSITY OF TURKU

Medium resolution (10 m-100 m) satellite data

+ Better suited for combination with field inventories
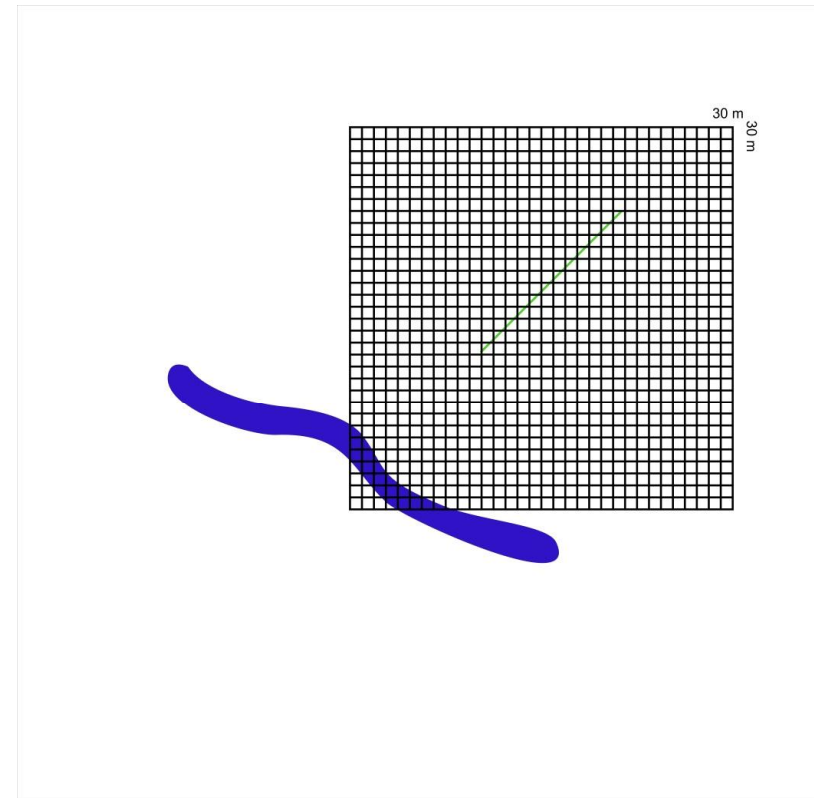
Issues when combining multiple scenes


Flight path

Flight path

1. Clouds / atmospheric contamination
2. Directional effects

UNIVERSITY OF TURKU

Taito workflow for image preprocessing and correction of directional effects (in R)
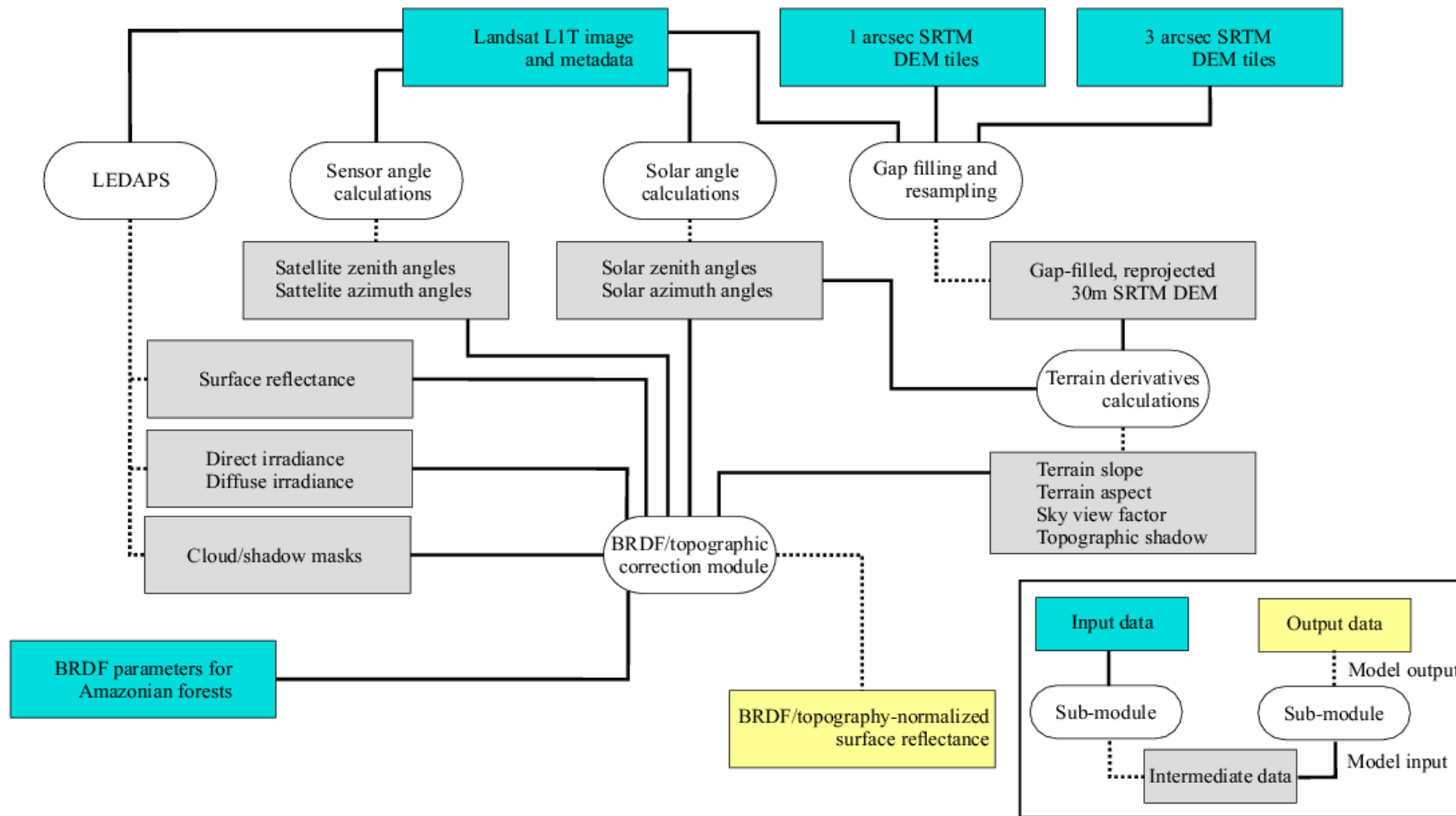
Image compositing

- Depending on selection criterion, improvement in radiometric consistency of composite image with increasing data

- Trade-off:
  - temporal resolution
 - radiometric consistency
  - computing resources

Selection criterion

UNIVERSITY OF TURKU
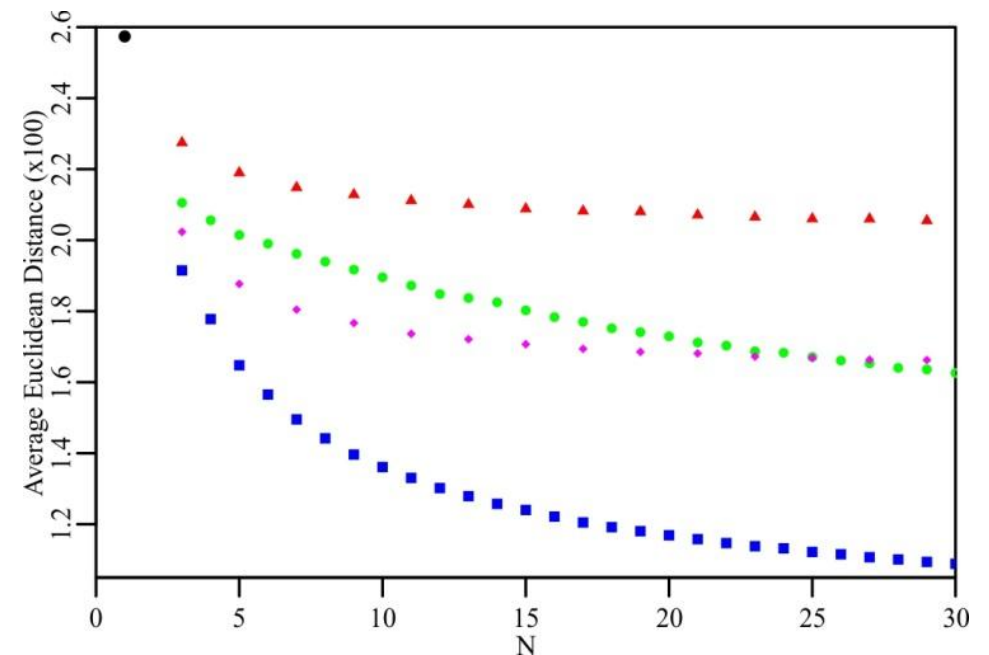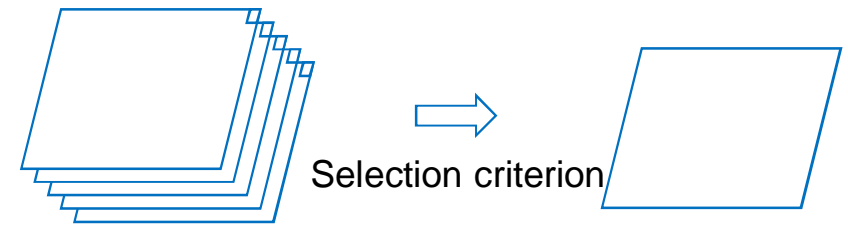
Application 1
Basin-wide vegetation modelling

- Combination with field surveys of species composition/soil

- Extremely small spectral differences between vegetation types

- Requires single composite image with high spatial consistency

- Compositing period: months July-September from 2000-2009

- ≈30.000 scenes (at ≈100MB per compressed scene)
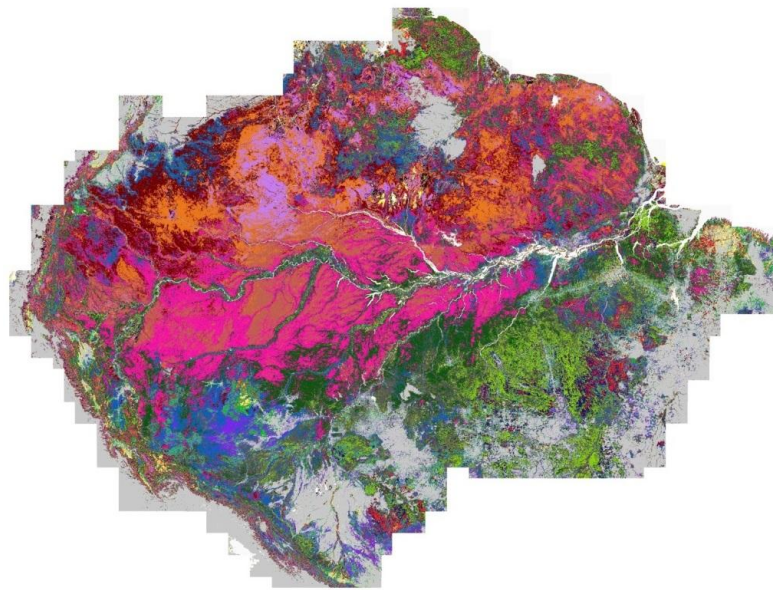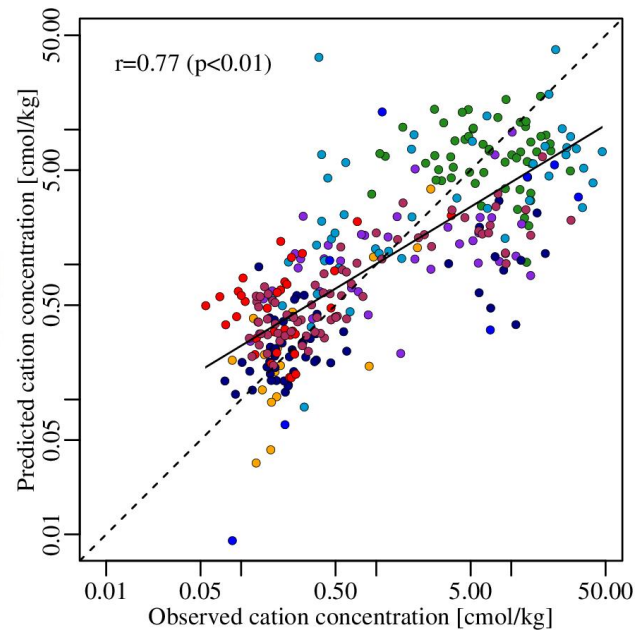

UNIVERSITY OF TURKU
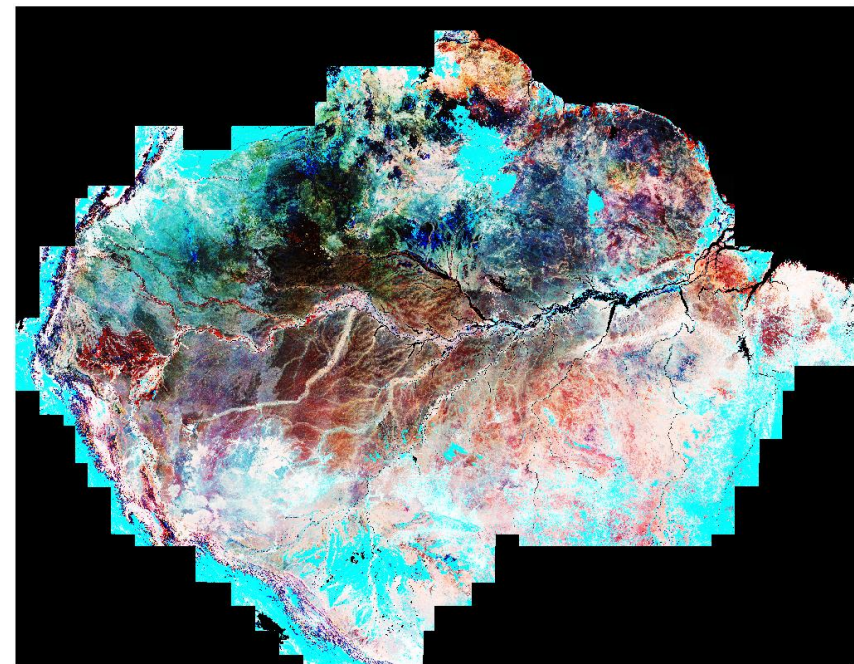
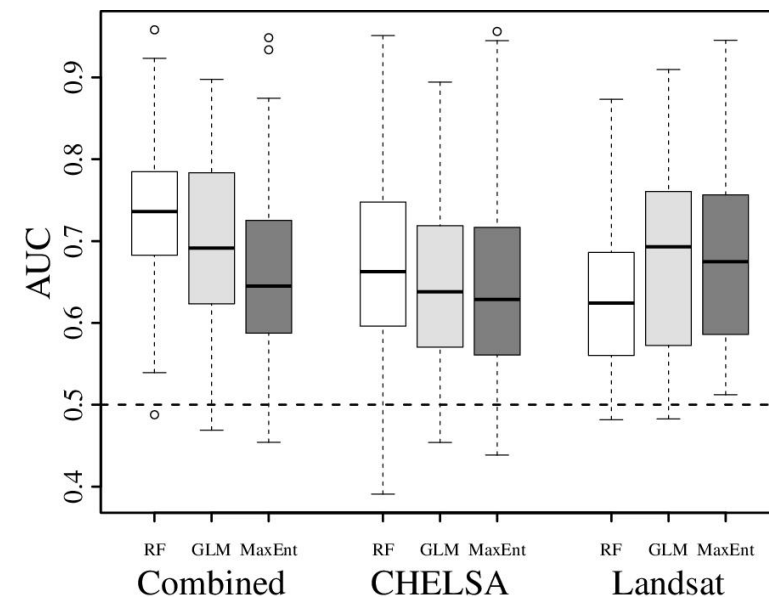# Application 1
Basin-wide vegetation modelling
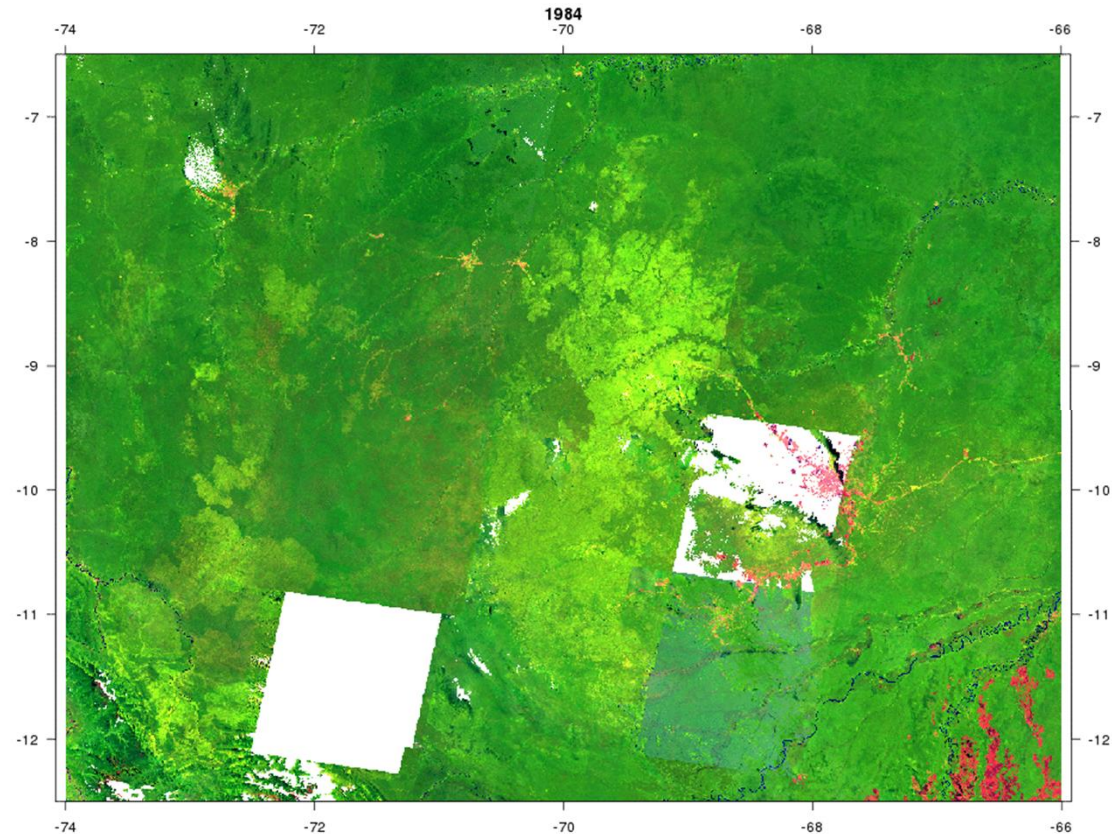
Unsupervised classification

Soil fertility modelling

Understory species distribution modelling

Application 2 Bamboo forests
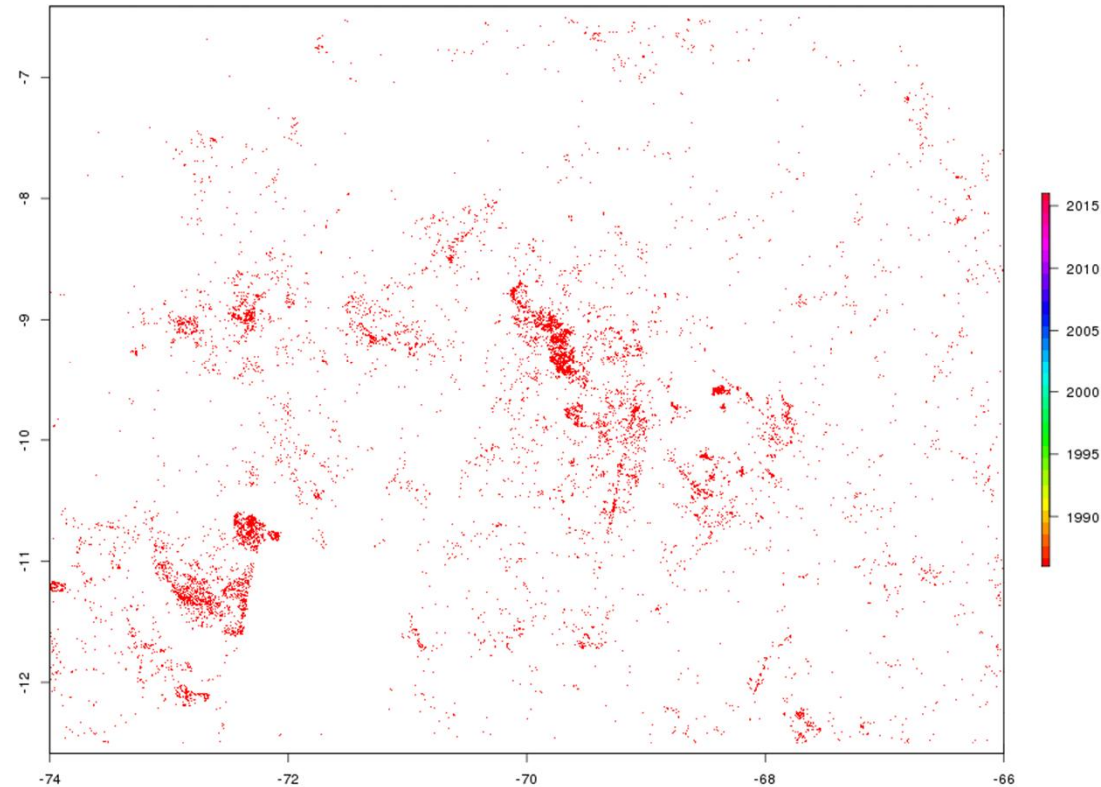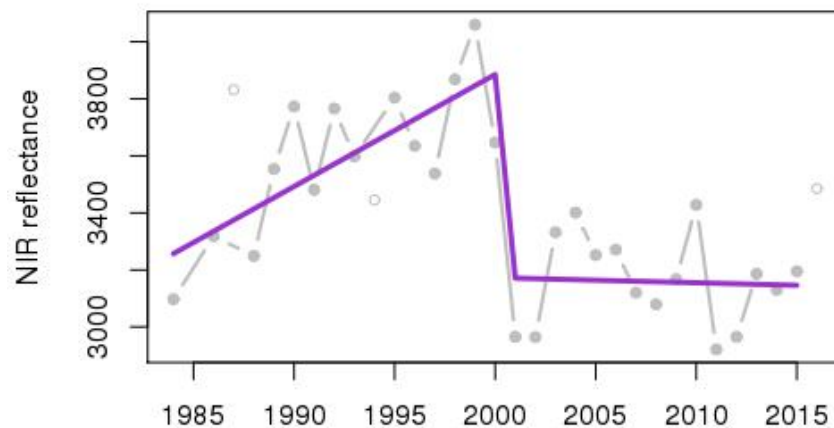
- Die-off/regrowth cycles of ≈27 years

- Synchronous patches up to order of 10 km diameter

- Yearly composites 1984-2017

Application 2
Bamboo forests

- Pixel-based time series analysis
- Dynamic time warping

Practical issues
Data volume

- Each 8-bit Landsat scene (.tar.gz) ≈ 100 MB
  - Raw data 30.000 scenes ≈ 3 TB
  - Processed imagery: 16-bit
- Landsat freely availably from USGS EarthExplorer
- For bulk data user, EarthExplorer Inventory service
  - JSON API in R
  - Querying and downloading integrated in processing chain
  - Allows only single download, no parallelization possible
  - Download speed max. ≈ 10 MB/sec (10 sec/scene) ==> 83 h.
- 30 m digital elevation model for Amazon region

UNIVERSITY OF TURKU

Practical issues
Data processing

- Processing of each raw Landsat scene
  - Cloud masking, atmospheric correction (top-of-atmosphere to surface reflectance), calculation of illumination/viewing angles, correction of topographic/directional effects
  - ≈ 1 h/scene ==> 1250 days
  - Straightforward parallelization per scene
- Combining individual Landsat scene in composite image
  - 30 m resolution for entire Amazon biome ≈ 100 GB
  - Not manageable for most personal computers
  - Tile-based approach, 0.5 degree tiles (to do: virtual raster)
  - For each tile, parallelization per image row

Practical issues
Data processing

- Time series analysis
  - See presentation Jan
  - Pixel-based
    - Data input/processing parallelized by data rows

Data distribution

- Work in progress

UNIVERSITY OF TURKU

Thank you!

jasper.vandoninck@utu.fi