# Change Detection in Spatiotemporal SAR Data for Deforestation Monitoring

*Johannes Hansen*

Doctor of Philosophy

The University of Edinburgh

July 9, 2022

# Declaration

I declare that this thesis has been composed solely by myself and that it has not been submitted, in whole or in part, in any previous application for a degree. Except where stated otherwise by reference or acknowledgment, the work presented is entirely my own.

Chapter 3 has been previously published in Hansen, Mitchard, et al. (2020), and chapter 6 has been published in Hansen (2022).

(*Johannes Hansen*)

# Abstract

Forests play a vital role in the wellbeing of our planet. Large and small scale deforestation across the globe is threatening the stability of our climate, forest biodiversity, and therefore the preservation of fragile ecosystems and our natural habitat as a whole. With increasing public interest in climate change issues and forest preservation, a large demand for carbon offsetting, carbon footprint ratings, and environmental impact assessments is emerging. Satellite remote sensing is the only method that can provide global coverage at frequent revisit times and is therefore the standard method for global forest monitoring. Most often, deforestation maps are created from optical data such as Landsat and MODIS. Although such maps are of generally good quality, they cannot quantify biomass and are not typically available at less than annual intervals due to persistent cloud cover in many parts of the world, especially the tropics where most of the world's forest biomass is concentrated. Synthetic Aperture Radar (SAR) can fill this gap as it penetrates clouds and interacts with the three-dimensional structure on the ground in a way that scales with volume and therefore biomass. While longer wavelengths are better for deeper penetration of the canopy and therefore biomass estimation, one of the most readily available data sources is Sentinel-1, a shorter wavelength C-band radar. In this thesis, the theory behind SAR is discussed to demonstrate its usefulness for forest monitoring. The potential for C-band SAR data to distinguish forest and non-forest is then assessed empirically in different regions of the world and existing change detection algorithms for deforestation are reviewed. The effect of spatial and temporal context on change detection accuracy is investigated, leading to the development of a robust method for deforestation detection in the absence of reliable reference data which often constitutes the largest practical hurdle. This method achieves a change detection sensitivity (producer's accuracy) above 99%, although false positives lead to a low user's accuracy of about 60%. The mean change detection delay amounts to about two to three months. While further work is required to reduce the false positive rate, improve detection delay, and validate this method in different biomes, the results show that Sentinel-1 data have the potential to advance global deforestation monitoring.

# Lay Summary

Forests play a vital role in the wellbeing of our planet. Large and small scale deforestation across the globe is threatening the stability of our climate, forest biodiversity, and therefore the preservation of fragile ecosystems and our natural habitat as a whole. With increasing public interest in climate change issues and forest preservation, a large demand for carbon offsetting, carbon footprint ratings, and environmental impact assessments is emerging. Satellite remote sensing is the only method that can provide global coverage at frequent intervals and is therefore the standard method for global forest monitoring. Most often, deforestation maps are created from optical images. Although such maps are of generally good quality, they cannot directly measure the amount of biomass and are not typically available at less than annual intervals due to persistent cloud cover in many parts of the world, especially the tropics where most of the world's forest biomass is concentrated. Radar images can fill this gap as they are unaffected by clouds. Radar is also reflected off the ground in a way that depends on the shape of the objects on the ground. Careful interpretation of these radar images therefore allows us to draw conclusions about the volume occupied by vegetation and in turn their biomass. Different radar satellites differ mainly in their wavelength, which dictates what size of objects they can observe. In general, forest biomass estimation works best with longer wavelengths. However, one of the most readily available radar data sources is the European Sentinel-1 satellite, which has a shorter wavelength radar instrument. In this thesis, the theory behind radar imaging for forest monitoring is discussed. The potential for Sentinel-1 data to distinguish forest and non-forest is then assessed for different regions of the world and existing methods for deforestation detection are reviewed. One of the biggest challenges in deforestation detection is often the lack of reliable reference data. For this reason, a robust method for deforestation detection in the absence of high quality reference data is proposed. This method achieves a high detection sensitivity although false positives lead to a low specificity. While further work is required to validate this method in different biomes and improve the deforestation detection result, including faster detection, the results show that Sentinel-1 has the potential to advance global deforestation monitoring.

# Acknowledgments

First and foremost, I would like to express my gratitude towards my supervisors Stuart King and Edward Mitchard, who have guided me through this endeavor and without whom this PhD would not have been possible.

I want to thank the Data Lab for funding this PhD, and my examiners for taking the time to read my thesis.

My sincere thanks go to the team at Ecometrica who have supported the creation of this PhD project and helped provide valuable direction in the early stages of the PhD, especially Ryan Elfman, Dimitrios Michelakis, and Richard Tipper.

Finally, I want to thank my family and friends, both in Edinburgh and around the world, for their unwavering support and for making my time here enjoyable. Thanks to those who have come to visit me in Edinburgh, those who couldn't, those who have been up for spontaneous trips when I needed a break, and those who have been fighting their very own battles. Thank you Feli for your patience, support, and joy. A final shoutout goes to Udhayan, Cinthia, Laura, and Tom who have all shared a living space with me at one point during my PhD and who have seen me at my best and at my worst. All of you have given my time here meaning and joy.

# Contents

# Glossary

**ALOS** *Advanced Land Observing Satellite*. 35, 94, 95

**BFAST** *Breaks for Additive Seasonal and Trend* (Verbesselt, Hyndman, Newnham, et al. 2010). 68, 69, 71, 72, 98, 99

**CCDC** *Continuous Change Detection and Classification*, a land cover change detection algorithm (Zhu and Woodcock 2014b). 68, 74

**CMFDA** *Continuous Monitoring of Forest Disturbance Algorithm* (Zhu, Woodcock, and Olofsson 2012). 72

**DCP** *Degree Confluence Project*, http://confluence.org/. 36, 63

**FAO** *Food and Agriculture Organization of the United Nations*. 3, 35

**Fmask** *Function of mask*, a cloud masking algorithm (Zhu and Woodcock 2012). 71, 72, 74, 95

**FRA** FAO *Global Forest Resources Assessment*, http://www.fao.org/forest-resources-assessment/en/. 36, 63

**GDAL** *Geospatial Data Abstraction Library*, https://www.gdal.org/. 143, 146

**GEDI** *Global Ecosystem Dynamics Investigation*, a NASA mission for forest measurement based on a space-borne full waveform lidar. 68

**LandTrendr** *Landsat-based Detection of Trends in Disturbance and Recovery*, a land cover change detection algorithm for Landsat (Kennedy, Yang, et al. 2010). 68, 74

**LEDAPS** *Landsat Ecosystem Disturbance Adaptive Processing System*, an atmospheric correction algorithm for Landsat (Schmidt et al. 2013). 71, 95

**lidar** *light detection and ranging*, a portmanteau of *light* and *radar*. 2, 68

**MODIS** *Moderate Resolution Imaging Spectroradiometer*, a NASA imaging sensor. 71

**MulTiFuse** *Multi-sensor Time-series correlation and Fusion* (Reiche, Verbesselt, et al. 2015). 95, 96, 98

**OLCI** *Ocean and Land Colour Instrument*, a medium-resolution imaging spectrometer on board the Sentinel-3 satellite. 3

**PALSAR** *Phased Array type L-band Synthetic Aperture Radar*, a SAR instrument on board the ALOS satellite. 35, 94, 95

**REDD+** *Reducing Emissions from Deforestation and forest Degradation and the role of conservation, sustainable management of forests and enhancement of forest carbon stocks in developing countries*. A framework by the UNFCCC. 2, 68

**SNAP** *Sentinel Application Platform*, a companion software for Sentinel data processing (http://step.esa.int/main/toolboxes/snap/). 25, 26

**SRTM** *Shuttle Radar Topography Mission*. 36, 42

**STL** *Seasonal-Trend Decomposition Procedure Based on Loess* (Cleveland et al. 1990). 70

**VCF** *Vegetation Continuous Fields*, a MODIS product. 71

# Acronyms

**AGB** Above Ground Biomass. 7, 8, 67

**API** Application Programming Interface. 25, 143

**BFM** BFAST-monitor. 96

**CLI** Command Line Interface. 143

**DCI** difference change index. 94

**DEM** Digital Elevation Model. 26, 33, 36, 42, 58, 59, 68

**DI** Disturbance Index (Healey et al. 2005; Masek et al. 2008). 73

**DTW** Dynamic Time Warping. 109–121, 123, 131

**ENL** Estimated Number of Looks. 37, 87

**EO** Earth Observation. 68

**EVI** Enhanced Vegetation Index. 68

**EW** Extra Wide Swath. 23

**F/NF** forest / non-forest. 97

**FLC** forest land cover. 94

**GCP** Ground Control Point. 27, 30–32, 159

**GHG** greenhouse gas. 68

**GLC** Global Land Cover. 63

**GRD** Ground Range Detected. 23–25, 63, 64, 160

**GV** Green Vegetation. 95

**IFZ** Integrated Forest Z-score (Huang et al. 2010). 73

**InSAR** Interferometric SAR. 8, 68

**IPCC** International Panel on Cimate Change. 1

**IW** Interferometric Wide Swath. 22–24, 42

**JM** Jeffreys-Matusita difference (Laliberte et al. 2012). 95, 97

**KDE** Kernel Density Estimation. 49

**kNN** k-Nearest-Neighbors. 38, 45, 53–55, 57, 62

**MOSUM** moving sum of residuals. 99

**MRV** Measurement, Reporting and Verification. 1

**MTL** Mean Time Lag. 97, 98

**NBR** Normalized Burn Ratio. 73

**NDFI** Normalized Difference Fraction Index (Souza, Roberts, et al. 2005). 7

**NDVI** Normalized Difference Vegetation Index. 2, 3, 68, 70, 73, 75, 95–100

**NFI** national forest inventory. 38

**NFMS** National Forest Monitoring System. 1

**NPV** non-photosynthetic vegetation. 95

**OA** Overall Accuracy. 69, 138, 140

**OLR** Ordinal Logistic Regression. 71

**OVL** Overlap Coefficient. 44

**PA** Producer’s Accuracy. 69, 109, 117, 131, 138, 140

**PDF** probability density function. 44, 96, 98

**PELT** Pruned Exact Linear Time (Killick et al. 2012). 92

**POD** Precise Orbit Determination. 42

**POEORB** Precise Orbit Ephemerides. 25

**PPI** Plant Phenology Index. 7

**PyPI** Python Package Index (https://pypi.org/). 145, 158

**QDA** Quadratic Discriminant Analysis. 45, 53, 55, 57, 63

**RESORB** Restituted Orbit. 25

**RF** Random Forest. 45, 53

**RGB** red, green, blue. 3

**RMSE** root mean square error. 108

**SAR** Synthetic Aperture Radar. 2, 3, 7, 8, 10, 11, 19, 33, 35, 37, 65, 68, 75–77, 94, 95, 98, 99, 141, 159, 160, 163, 165

**SLC** Single Look Complex. 18, 23–25, 42, 62, 64, 77, 159, 160

**SM** Stripmap. 22

**SMA** Spectral Mixture Analysis. 94, 95

**SOTR** Space Objects Tracking Radar. 8

**SPI** Standardized Precipitation Index. 98, 99

**SVM** Support Vector Machine. 45, 53–55, 57, 62, 64, 65

**SWIR** Shortwave Infrared. 73

**TLS** Terrestrial Laser Scanning. 7

**TOA** Top Of Atmosphere. 73

**TOPSAR** Terrain Observation with Progressive Scans SAR. 22, 26

**TRMM** Tropical Rainfall Measurement Mission. 98

**UA** User’s Accuracy. 69, 109, 117, 131, 138, 140, 163

**UNFCCC** United Nations Framework Convention on Climate Change. 1

**USGS** United States Geological Survey. 42

**WV** Wave mode. 23

# Chapter 1

# Introduction

Forests play a major role in our fragile ecosystem. When planted they act as a carbon sink, but when cut down or burnt they act as a carbon source. In light of the universal awareness of the reality and consequences of global climate change, significant international efforts are being undertaken to preserve forests and reforest areas that have experienced large scale deforestation. Examples of these undertakings include the Bonn Challenge, the New York Declaration on Forests, and the Paris Agreement. The first step for prevention of deforestation must be accurate monitoring and mapping of affected areas with as little delay as possible. REDD+ (which stands for “Reducing emissions from deforestation and forest degradation and the role of conservation, sustainable management of forests and enhancement of forest carbon stocks in developing countries”) was adopted by the United Nations Framework Convention on Climate Change (UNFCCC) to reverse deforestation and forest degradation, and hence their contribution to climate change. It requires developing countries to implement Measurement, Reporting and Verification (MRV) systems as part of a National Forest Monitoring System (NFMS), such as the Brazilian DEGRAD system (INPE n.d.). For REDD+ to succeed it is vital that monitoring systems are accurate, scalable, trusted, and comparable. The International Panel on Cimate Change (IPCC) defines quality standards for these measurement systems (IPCC 2003), and is specific that remote sensing should be used, but no specifics about the precise technologies and methods are provided. Similarly, all countries require good maps of their changes in forest areas as part of their reporting on changes in all carbon pools, related to their territory-wide reporting to the UNFCCC on carbon sinks and sources and how these compare to their commitments under the Paris Agreement.

Due to large-scale coverage and regular revisit times, satellite remote sensing is the only viable candidate for large-scale forest cover monitoring systems. Frequent revisits are necessary not only for precise dating of deforestation events, but also because of rapid regrowth, which may make clearings indistinguishable from intact forest within less than two years (Herold and Skutsch 2011). Data available from satellite sensors encompass optical imagery (including derivatives such as the Normalized Difference Vegetation Index (NDVI)), Synthetic Aperture Radar (SAR), and lidar data, each with different advantages and drawbacks. A big challenge for remotely sensed forest monitoring is the quantification of forest degradation in addition to deforestation (Herold, Román-Cuesta, et al. 2011), as it constitutes more subtle changes such as extraction of fuel wood and understory grazing (Skutsch et al. 2011) in which the forest canopy may remain intact. For such cases, field measurements remain an important source of reference data. Many authors have suggested the use of satellite data for mapping forest area and forest properties for such national monitoring purposes. For example, McRoberts et al. (2007) elaborate on the potential of remote sensing data for national forest inventories. They emphasize the low cost of measuring areal quantities such as forest area or volume compared to conventional field sampling. However, they note limitations in quantifying the uncertainty of forest inventories produced in this way.

Unsustainable logging is the main cause of forest degradation in tropical and subtropical areas (Hosonuma et al. 2012), largely due to commercial timber demand. Further causes include over-grazing, fire, and charcoal production. The tropics are also the area with the largest total forest loss, with South America exhibiting particularly high deforestation rates (Hansen, Potapov, et al. 2013).

Reiche, Lucas, et al. (2016) published a commentary in which they note that the majority of REDD+ member countries in the tropics are using Landsat, but no SAR data, for their respective national monitoring systems. That is even though most of these countries are subject to persistent cloud cover throughout the year, resulting in large data gaps for any optical sensor; a problem which demonstrably does not affect SAR sensors at C-band or longer (Lu 2006). Scientists have already successfully demonstrated the capabilities of SAR for deforestation monitoring (Bouvet et al. 2018; Mitchard, Saatchi, Lewis, et al. 2011; Rahman et al. 2010; Soja, Persson, et al. 2018) and Lehmann, Caccetta, et al. (2015) have presented a strategy for combining SAR and optical data in a large-scale forest monitoring system. Reiche, Lucas, et al. attribute this absence of SAR data in national monitoring systems to the comparative ease of processing optical data, the larger historic archives, and the limited availability of free SAR data and SAR processing tools (prior to the launch of Sentinel-1A almost all satellite SAR data were available only on commercial terms). As most global forest maps have been derived from optical imagery—requiring a mosaic over multiple time steps for a cloud-free composite—they have limited temporal repeatability. Two notable exceptions of repeated global forest maps are the maps by Hansen (Hansen, Potapov, et al. 2013) and the FAO (Lindquist et al. 2012), though in practice it is likely that some deforestation is missed or reported late due to cloud cover in these studies too. Moreover, while the Hansen map (Hansen, Potapov, et al. 2013) provides continuous cover, the FAO map only samples select spatially distributed boxes (Hansen, Potapov, et al. 2013), meaning it provides useful regional statistics but cannot be used to locate actual deforestation events.

## 1.1 Measuring Forest

How do we measure forest from space?

This question can be more broadly framed as measuring vegetation in general. The resolution of most satellite imagery is too coarse to discern individual trees or other plants. We are therefore left with a single pixel that may or may not contain forest, other types of vegetation, as well as other natural or artificial objects. This pixel has an intensity value for each wavelength of the sensor (i.e., a *color*). The intensity of the full pixel represents a mixture of the signatures of all reflecting objects within that pixel. The number of wavelengths may be three if we take a simple photograph in red, green, blue (RGB), or (usually) more, in which case we speak of *multispectral* data. To give some examples, Landsat measures 9 spectral bands, Sentinel-2 measures 13, and Sentinel-3 OLCI as many as 21 spectral bands. Going even further, *hyperspectral* data may contain hundreds of bands at very narrow bandwidths.

In the past, most remote sensing approaches to assessing plant phenology, whether in a single image or over time, reduced the available multispectral data to a single vegetation index. The basic idea is to capture the essence of “What makes vegetation?” in a single quantity. A large number of such indices have been proposed, the most well-known of which is arguably the NDVI. A non-exhaustive list is given here:

**NDVI** Normalized Difference Vegetation Index (Rouse et al. 1974)

$$\mathrm{NDVI} = \frac{\text{NIR} - \text{red}}{\text{NIR} + \text{red}}$$

**DVI** Difference Vegetation Index (Tucker 1979)

$$\mathrm{DVI} = \text{NIR} - \text{red}$$

**SAVI** Soil Adjusted Vegetation Index (Huete 1988)

$$\mathrm{SAVI} = \frac{\text{NIR} - \text{red}}{\text{NIR} + \text{red} + L}(1 + L)$$

where $L$ is a correction factor which ranges from 0 for very high vegetation cover to 1 for very low vegetation cover. The most typically used value is 0.5 which is for intermediate vegetation cover.

**TSAVI** Transformed Soil Adjusted Vegetation Index (Baret, Guyot, and Major 1989)

$$\mathrm{TSAVI} = \frac{a(\text{NIR} - a \cdot \text{red} - b)}{a \cdot \text{NIR} + \text{red} - ab + X(1 + a^2)}$$

where $a$ is the slope of the soil line, $b$ is the intercept of the soil line, and $X = 0.8$ is an adjustment factor. The median soil line values reported in Baret and Guyot (1991) are $a = 1.2$ and $b = 0.04$.

**GEMI** Global Environmental Monitoring Index (Pinty et al. 1992)

$$\mathrm{GEMI} = \eta(1 - 0.25\eta) - \frac{\text{red} - 0.125}{1 - \text{red}}$$

where

$$\eta = \frac{2 * (\text{NIR}^2 - \text{red}^2) + 1.5\text{NIR} + 0.5\text{red}}{\text{NIR} + \text{red} + 0.5}$$

**GNDVI** Green Normalized Difference Vegetation Index (Buschmann et al. 1993)

$$\mathrm{GNDVI} = \frac{\text{NIR} - \text{green}}{\text{NIR} + \text{green}}$$

**EVI** Enhanced Vegetation Index (Huete, Justice, et al. 1994)

$$\mathrm{EVI} = G \cdot \frac{\text{NIR} - \text{red}}{\text{NIR} + C_1 \cdot \text{red} - C_2 \cdot \text{blue} + L}$$

where $G = 2.5, C_1 = 6, C_2 = 7.5, L = 1.$

**MSAVI2** Modified Soil Adjusted Vegetation Spectral Index (Qi, Chehbouni, et al. 1994)

$$\mathrm{MSAVI2} = \frac{2\text{NIR} + 1 - \sqrt{(2\text{NIR} + 1)^2 - 8(\text{NIR} - \text{red})}}{2}$$

**OSAVI** Optimized Soil Adjusted Vegetation Index (Rondeaux et al. 1996)

$$\mathrm{OSAVI} = \frac{\text{NIR} - \text{red}}{\text{NIR} + \text{red} + L}(1 + L)$$

**GARI** Green Atmospherically Resilient Index (Gitelson, Kaufman, and Merzlyak 1996)

$$\mathrm{GARI} = \frac{\text{NIR} - [\text{green} - \lambda(\text{blue} - \text{red})]}{\text{NIR} + [\text{green} - \lambda(\text{blue} - \text{red})]}$$

where $\lambda$ is a parameter controlling the atmospheric correction.

**VI$_{\text{green}}$** Vegetation Index (green) (Gitelson, Kaufman, Stark, et al. 2002)

$$\mathrm{VI}_{\mathrm{green}} = \frac{\text{green} - \text{red}}{\text{green} + \text{red}}$$

**VARI$_{\text{green}}$** Visible Atmospherically Resistant Index (green) (Gitelson, Kaufman, Stark, et al. 2002)

$$\mathrm{VARI}_{\mathrm{green}} = \frac{\text{green} - \text{red}}{\text{green} + \text{red} - \text{blue}}$$

**GDVI** Green Difference Vegetation Index (Sripada et al. 2006)

$$\mathrm{GDVI} = \text{NIR} - \text{green}$$

**NG** Normalized Green (Sripada et al. 2006)

$$\text{NG} = \frac{\text{green}}{\text{NIR} + \text{red} + \text{green}}$$

**NR** Normalized Red (Sripada et al. 2006)

$$\text{NR} = \frac{\text{red}}{\text{NIR} + \text{red} + \text{green}}$$

**NNIR** Normalized Near Infrared (Sripada et al. 2006)

$$\text{NNIR} = \frac{\text{green}}{\text{NIR} + \text{red} + \text{green}}$$

**GRVI** Green Ratio Vegetation Index (Sripada et al. 2006)

$$\text{GRVI} = \frac{\text{NIR}}{\text{green}}$$

**GSAVI** Green Soil Adjusted Vegetation Index (Sripada et al. 2006)

$$\text{GSAVI} = \frac{\text{NIR} - \text{green}}{\text{NIR} + \text{green} + L}(1 + L)$$

where $L = 0.5$ (see SAVI).

This is just a subset of available indices, and many more are bound to be developed. In the past, these vegetation indices were a valuable asset to the remote sensing community as they provided a simple metric to assess vegetation. They are easy to compute, analyze, visualize, and interpret. However, as the long list of vegetation indices clearly shows, a single index cannot represent the entirety of information contained in multispectral data. Especially considering the increasing necessity to fuse multiple data sources, condensing all information into a single index appears outdated in the era of high performance computing and machine learning.

Furthermore, all of these vegetation indices are derived from physical principles such as known absorption spectra. While commendable from a scientific perspective, such an approach becomes unfeasible in the context of big data, potentially hundreds of different kinds of vegetation and non-vegetated surfaces that we would all like to be able to distinguish. Manually crafting indices for each of these from tens of variables based on prior knowledge is impossible. That is why machine learning has become so popular also in remote sensing, because it attempts to make sense of large quantities of data without relying on physical modeling.

Balzter, Cole, et al. (2015) demonstrated the capability of Sentinel-1 data for land cover classification, when used in conjunction with topographic and geomorphometric datasets. Seasonality may also be an important factor in distinguishing between different surface covers or species. Zhang, Friedl, et al. (2009) assessed the ability to correctly detect vegetation phenology depending on the temporal resolution of the available data, as well as the impact of missing data. They found that temporal resolutions of 6–16 days performed optimally in determining the phenology, whereas both smaller and larger resolutions introduced errors. The metrics measured were the onset dates of the phenological phases (i) greenup, (ii) maturity, (iii) senescence, and (iv) dormancy. By fitting piecewise logistic models to the data, the authors were able to demonstrate that for temporal resolutions of $\leq 16$ days, the absolute error in these metrics was less than 3 days. The relative phase shift of the sampling (i.e., at what dates exactly the data were available) had no impact on this accuracy. In a related attempt at measuring and quantifying plant phenology from remote sensing data, Jin et al. (2014) have developed the definition of a Plant Phenology Index (PPI).

Other indices have been developed to directly quantify the degree of forest damage, such as the Normalized Difference Fraction Index (NDFI) (Souza, Roberts, et al. 2005).

When measuring forest, the gold standard is to be able to not just classify forest and non-forest and thus detect deforestation, but to measure the Above Ground Biomass (AGB) content in the forest. The increasing public interest in climate change issues and forest preservation, along with emerging climate protection legislation, has created a large market for carbon offsetting, carbon footprint ratings, and environmental impact assessments, all of which require accurate estimates of the forest biomass. Quantification of biomass content is also necessary to detect forest degradation, which generally constitutes a reduction in forest biomass, while still leaving the area forested. According to some sources, forest degradation now accounts for more loss of biomass than deforestation (Qin et al. 2021). The correlation between backscatter intensity and AGB has been investigated in numerous studies (Mitchard, Saatchi, Lewis, et al. 2011; Mitchard, Saatchi, Woodhouse, et al. 2009) and is generally contentious (Mitchard, Feldpausch, et al. 2014; Woodhouse et al. 2012). Allometric relationships are derived from AGB ground truth data that are either obtained from destructive sampling (Chave et al. 2005) or Terrestrial Laser Scanning (TLS) (Calders, Adams, et al. 2020; Calders, Newnham, et al. 2015; Gonzalez de Tanago et al. 2018; Momo Takoudjou et al. 2018). There is a better overall correlation for longer wavelengths such as L-band (Lucas, Bunting, et al. 2010; Luckman et al. 1998; Mermoz and Le Toan 2016; Mitchard, Saatchi, Lewis, et al. 2011) compared to shorter wavelengths like the Sentinel-1 C-band due to faster backscatter saturation at shorter wavelengths (Balzter 2001; Dobson et al. 1992; Lucas, Milne, et al. 2000). Saturation of allometric relationships between AGB and remotely sensed data generally starts between $15$ to $70\,\mathrm{t\,ha^{-1}}$ for optical-derived indices (Lu 2006), $20$ to $30\,\mathrm{t\,ha^{-1}}$ for C-band SAR (Lucas, Milne, et al. 2000), $40$ to $150\,\mathrm{t\,ha^{-1}}$ for L-band SAR (Balzter, Talmon, et al. 2002; Lu 2006; Mitchard, Saatchi, Woodhouse, et al. 2009), and $80$ to $200\,\mathrm{t\,ha^{-1}}$ for P-band SAR

(Dobson et al. 1992; Lucas, Milne, et al. 2000). This saturation leads to large uncertainties in global biomass maps (Mitchard, Saatchi, Baccini, et al. 2013).

Several studies have used Sentinel-1 for AGB mapping, but generally only in conjunction with other data sources such as Sentinel-1 and Sentinel-2 (Chen et al. 2018; Forkuor et al. 2020; Nuthammachot et al. 2020), or Sentinel-1 and Landsat-8 (Li et al. 2020). Combining multiple sources is a very common approach for estimating the AGB: other studies have combined lidar and SAR (Kaasalainen et al. 2015; Mitchard, Saatchi, White, et al. 2012), multi-wavelength SAR (C-band, L-band) and multispectral optical (Laurin et al. 2018), or optical, SAR, and lidar (Zimbres, Rodríguez-Veiga, et al. 2021). Some have even used Interferometric SAR (InSAR) (Balzter 2001) or the Sentinel-1 temporal coherence (Cartus, Santoro, Wegmuller, et al. 2021). In general, the accuracy of AGB measurements is increased when using time series data (Santoro et al. 2021). As longer wavelengths are generally preferable for biomass estimation, ESA is launching a novel P-band satellite mission appropriately named BIOMASS (Quegan, Le Toan, et al. 2019) that will be uniquely suited for AGB estimation in tropical forests (Soja, Quegan, et al. 2021). It was originally planned to launch in 2022, and is currently scheduled for 2023. It will be the first P-band SAR instrument in space and will provide global AGB and tree height estimates with the exception of Europe and North America due to US DoD Space Objects Tracking Radar (SOTR) restrictions.

Another forest parameter that is often measured is tree height, which can also serve as a proxy for the AGB (Balzter, Rowland, et al. 2007). It is most easily obtained from lidar data (Kwak et al. 2007; Nilsson 1996; Verhelst et al. 2021) and local forest height data can be used to calibrate and improve AGB and tree height estimates derived from SAR (García et al. 2018; Pourshamsi, Garcia, et al. 2018; Pourshamsi, Xia, et al. 2021). In particular, polarimetric and interferometric SAR are often used for tree height estimation (Balzter, Luckman, et al. 2007)

These considerations lead me to define the basic research question, or rather, research purpose. The goal is to develop a robust detection method for deforestation and forest degradation from multivariate and multi-source data. The method should have a number of qualities: (1) It should be physics-agnostic to a reasonable degree, such that prior knowledge about the input data may improve the result, but is not required. (2) It should provide reliable deforestation alerts even in the presence of sub-optimal reference data. (3) It should scale well with increasing dataset size.

## 1.2 Reference Data

One of the biggest challenges in deforestation monitoring is the (un)availability of good reference data. Many detection algorithms rely on some form of supervised machine learning, and any such algorithm is only as good as its training data. The difficulty in acquiring these reference data lies in the global scale of the phenomenon, its temporal sensitivity (many detection methods may fail even a few weeks after deforestation as some vegetation has regrown), local political tension including armed conflicts, the remote location of some affected areas, and many more (Rodríguez-Veiga, Wheeler, et al. 2017).

In practice, most reference data are generated by manually drawing polygons around areas that *look* like deforestation in high-resolution aerial or satellite imagery. Generally, there is a trade-off between spatial and temporal cover for any reference data set. Temporally dense data are only available over very few selected monitoring sites that may or may not be representative of forest elsewhere. Reference data of forest degradation (such as selective logging) are even more difficult to obtain.

The most widely used global data set is published at *GlobalForestWatch* (http://globalforestwatch.org/). In a highly cited paper, Hansen, Potapov, et al. (2013) present findings from a global survey of annual forest cover, gain, and loss for the years 2000 to 2012. Their analysis is based on Landsat data, resulting in a global spatial resolution of $30\,\mathrm{m}$. The highest rates of deforestation were found in the tropical regions, this being the only region to exhibit a clear trend of increasing annual deforestation. Indonesia was specifically noted to have the largest increase in yearly forest loss. The authors point to Brazil as a notable exception, with an actual decrease in deforestation and spatially explicit documentation of the forest extent. The most significant causes of forest dynamics and loss were determined to be forestry in temperate forests, and fires in boreal forests. Subtropical regions had the largest forest change dynamic (combined loss and gain) due to extensive forestry.

However, these data are not suitable for use as training data for machine learning algorithms, as they are themselves the output of a classification algorithm. The lack and quality of reference data remains an issue to be solved for deforestation monitoring. Table 1.1 summarizes some of the available reference data sets for deforestation detection.

**Table 1.1:** This table shows some available reference data sets.

| Name | Frequency | Period | Coverage | Download |
|---|---|---|---|---|
| GlobalForestWatch | annual | 2000 – 2020 | global | [a] |
| JRC FOROBS | annual | 1985 – 2020 | global | [b] |
| Intact Forest Landscapes (IFL) | 2000, 2013, 2016 | | global | [c] |
| DETER/PRODES | | 2008 – 2020 | Brazil | [d] |
| MapBiomas | annual | 1985 – 2019 | Brazil | [e] |

---

[a] https://data.globalforestwatch.org/documents/134f92e59f344549947a3eade9d80783/explore
[b] https://forobs.jrc.ec.europa.eu/TMF/download/
[c] http://www.intactforests.org/data.ifl.html
[d] http://terrabrasilis.dpi.inpe.br/en/download-2/
[e] https://mapbiomas.org/en/colecoes-mapbiomas-1?cama_set_language=en

## 1.3 Synopsis

This thesis is structured by chapters to gradually build up the concepts behind SAR from its foundation in physics to more applied topics like sensor details and data calibration. Following this, I go on to assess the potential of SAR data to distinguish forest and non-forest before finally introducing SAR-based change detection algorithms.

Chapter 2 is dedicated to describing the differences between optical and SAR imaging and explores the theory behind radar imaging, including concepts such as polarization, range and azimuth resolution, synthetic aperture, coherence, and speckle. I further introduce the Sentinel-1 data including platform specific terminology in the context of the theory outlined before. I then go on to explain the various pre-processing steps necessary to turn the SAR products as obtained from the data dissemination server into *analysis ready data*. In addition, this chapter contains an overview of the concepts of geocoding and data reprojection, including an analysis of the warping errors introduced in this process.

Having covered the basics of SAR processing, chapter 3 looks at the empirical separability of forest and non-forest using C-band SAR data. Detecting changes in the forest cover using SAR data requires a statistically significant difference in the pixel measurement of some radar parameter (e.g., backscatter, coherence, interferometric height, or texture) that can be reliably used to distinguish forested and non-forested area. At its core this is a classification problem as we are trying to assign a class (forest or non-forest) to each pixel given a series of measurements. This chapter therefore includes a literature review of attempts to classify forest and non-forest based on SAR data. I then assess empirically how well forested areas can be distinguished using dual-polarimetric C-band backscatter data, the easiest SAR data to use. The assessment is done for a variety of forest types by looking at forest sites in Alaska, Colombia, Finland, Florida, Indonesia, and the UK.

Chapter 4 introduces the core of this thesis. I review the literature on change detection algorithms and present results from the implementation of an existing algorithm. Many of these algorithms follow similar patterns so for my review I have grouped them together in two general categories; *model-based* and *non-parametric* algorithms. As most of these change detection algorithms operate on the time series of a single pixel, they often neglect the information provided by the spatial neighborhood of the pixel. One section of this chapter is therefore dedicated to addressing the information content of the spatial context, i.e. whether and how pixelwise change detection can be improved by considering neighboring pixels. As both SAR and optical data have their advantages and disadvantages, I include an additional section in which I review methods that aim to combine multiple data sources for improved change detection results.

In chapter 5, I address some of the main challenges in deforestation detection that have emerged throughout my analysis in the previous chapter. I present a novel change detection method that operates based on the distance to a reference forest class in a semi-supervised approach.

Finally, in chapter 6 I observe a lack of satisfactory Python libraries for EO data, especially for multivariate and multidimensional data. I therefore propose the creation of a unifying Python framework to facilitate algorithm development.

I conclude with a summary of the central findings from my thesis in chapter 8. I embed my results in the wider context of deforestation detection research and offer a discussion of the main observations. This leads me to formulate a number of open research questions that are left to be addressed in future work.

# Chapter 2

# Radar Theory

SAR imaging differs from optical imaging in many ways. Some of the conceptual differences result from the fact that radar imaging is (usually) an *active* imaging technique, while optical imaging is *passive*. This terminology refers to the fact that the radar system itself is the source of illumination for the target, while an optical camera measures the reflected portion of the light emitted from the sun (or other sources of light, such as lava, fires or artificial lighting).

Crucially, this means that for a radar system the incidence angle of the radiation is always equal to the look angle. As a consequence of this, an active radar imaging system is never imaging at nadir (nadir being the direction pointing straight down, perpendicular to the Earth's surface), as that would result in a hot spot in the center of the image due to strong specular reflection.

Whereas optical sensors measure the reflectance for each wavelength as a function of the look angle, radar measures the reflectance (for radar this is called *backscatter*) as a function of *range*. No explicit information about the direction of the return signal is contained in the received backscatter signal and it could have been scattered from anywhere within the field of view. However, if the three-dimensional location of the radar platform, the look angle, and the topography of the Earth surface are precisely known, the return time of the signal is sufficient to locate its origin on the ground. This also implies that there is no inherent concept of *pixels* in the radar antenna. Any radar image obtained as a raster of pixels is never the raw data, but a carefully crafted representation of the raw data using considerable mathematics to put each pixel in its place. A small part of the mathematics behind that process is covered in this chapter.

## 2.1 Radar Equation

One of the central equations governing active microwave systems is the radar equation, relating the received power $P_r$ to the transmitted power $P_t$:

$$P_r = P_t \frac{\sigma A_e G}{(4\pi R^2)^2} \tag{2.1}$$

where $\sigma$ is the radar cross-section of the target, $A_e$ is the effective area (or cross-section) of the antenna, $G$ is the gain of the antenna, and $R$ is the distance between the antenna and the target. The gain describes the ability of the antenna to focus the transmitted energy into a narrow beam. In practice, the gain and effective area are related by the following equation:

$$A_e = \frac{G\lambda^2}{4\pi} \tag{2.2}$$

where $\lambda$ is the wavelength such that the radar equation may be written as

$$P_r = P_t \frac{\sigma G^2 \lambda^2}{(4\pi)^3 R^4} \tag{2.3}$$

Because the transmitted power is distributed equally over the surface of a sphere with the source at its center (or a fraction of that surface for a focused beam), the power arriving at the target decays with the square of the distance. The power diminishes by a further quadratic factor because the scattered signal has to arrive back to the antenna. As a result the antenna has to be able to emit radiation in the order of Megawatts while detecting signals in the order of nanowatts.

$P_t$ and $G$ are properties of the antenna, while $P_r$ is the quantity measured. $R$ can be obtained from the signal return time, leaving the target cross-section $\sigma$ to be calculated. Dividing this cross-section by the area it covers gives the normalized radar cross-section:

$$\sigma_0 = \frac{\sigma}{A} \tag{2.4}$$

This quantity is also known as *backscatter*.

## 2.2 Range and Topography Effects

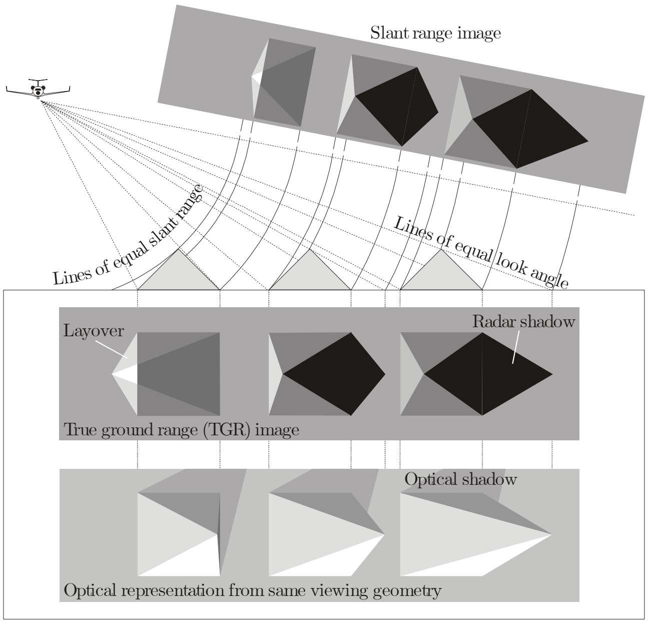


**Figure 2.1:** This figure illustrates key differences between the geometry of an optical camera and a radar imaging system. Graphic taken from Woodhouse (2005).

The term *range* is ambiguous, which is why in radar terminology we distinguish between *slant range* and *ground range*. Slant range refers to the axis connecting the antenna with the target and is directly proportional to the signal return time. It is the coordinate system in which the radar antenna operates. As the name implies, ground range refers to the distance on the ground, e.g. measured from the sub-satellite track. This is the coordinate system that we are used to and that we reasonably expect data products to obey. For a completely flat surface, the conversion from slant range to ground range follows a trivial trigonometric relation. However, in reality the ground is never completely flat, causing a series of artifacts in a radar image, as described in the following.

Two effects resulting from the non-zero elevation of targets on the ground are *layover* and *shadow*, both of which typically occur together. Because of the geometry of radar imaging, elevated objects appear to be leaning towards the imaging system, as opposed to leaning away from it in optical imaging. As a consequence, the radiation reflected from the nearside of the object is spread over a smaller angular region such that the measured amplitude per pixel is higher. The opposite is true for the farside of the object such that it will be appear darker. In some cases, the geometry will be such that the return from the top of an object appears closer than its base, even though in ground range it is further away, as illustrated in the leftmost case in fig. 2.1. This is known as the layover effect. Radar shadow occurs where the elevation of an object obscures objects behind it, such that no backscatter may be returned from a region of the image. That region is then imaged with backscatter values of exactly zero (this effect is shown in the middle and rightmost case in fig. 2.1).

These effects are inherent features of radar imaging rather than calibration errors as even an accurate topography model will not correct for them. There is no equivalent of these effects in optical imaging. The conceptual difference between radar shadow and optical shadow is also shown in fig. 2.1.

By default, topography is ignored and the image is projected onto a flat Earth. This results in errors which need to be corrected as part of a series of calibration steps (see *Terrain Correction* in section 2.8.2).

## 2.3 Synthetic Aperture Radar

So far we have only considered a one-dimensional image, because it is easy to visualize and explain. But in reality, the Earth surface is two-dimensional. If we are trying to map a return signal onto the ground based on the range calculated from the return time, we thus end up with a circle of possible sources rather than a single point. How do we tell where on this circle the target lies? In essence, the resolution in azimuth direction is determined by the beam width on the ground. How small the beam can get is subject to the diffraction limit just like for optical sensors. The angular bandwidth (based on a $3$ dB drop-off in amplitude), also known as *field of view*, depends on the wavelength $\lambda$ and the antenna size $d$. It can be approximated as

$$\theta_{3dB} \approx \frac{\lambda}{d} \tag{2.5}$$

The azimuth resolution is then given by

$$\rho_a = \frac{\lambda}{d} R \tag{2.6}$$

with an additional dependency on the distance to the ground, $R$. For Sentinel-1 we have $\lambda \approx 5.5$ cm, $d = 12.3$ m, and $R \approx 700$ km (at nadir), which gives a resolution of about $3$ km. Clearly, we hope to achieve a better result than that.

The solution is to synthesize a larger antenna, moving from *real aperture* to *synthetic aperture* radar. The basic idea is that if we have a wide beam on the ground, a single point on the ground will be within the satellite's field of view for as long as it takes the satellite to traverse that distance. Combining the data from all azimuth angles thus yields a much larger effective antenna size. Without going into detail on how this synthesis is performed, the algorithm is further exploiting the fact that the return signal is Doppler shifted due to the movement of the satellite. More specifically, the wave is blue-shifted in the direction of travel and red-shifted in the opposite direction. This effect is crucial for disentangling signals from different azimuth angles.

Having established the azimuth resolution, we can also investigate the resolution in the range direction. Range resolution is achieved by emitting radar waves in a series of short pulses. The problem of distinguishing between targets of similar range thus translates into distinguishing between two pulses of a small time delay. The (slant) range resolution $\rho_r$ is given by

$$\rho_r = \frac{c\tau_p}{2} \tag{2.7}$$

where $c$ is the speed of light and $\tau_p$ is the pulse length.

To improve the range resolution without decreasing the pulse length, radar systems apply a frequency modulation to the pulse, also called *chirping*. A chirped pulse with bandwidth $\Delta\nu$ has an *effective pulse length* of $\tau_e = 1/\Delta\nu$, which means the range resolution changes to

$$\rho_r = \frac{c}{2\Delta\nu} \tag{2.8}$$

The *ground range resolution* is then given by

$$\rho_g = \frac{\rho_r}{\sin\theta_i} \tag{2.9}$$

where $\theta_i$ is the incidence angle. This relation implies that the ground range resolution improves (i.e. $\rho_g$ decreases) as the incidence angle $\theta_i$ increases, which may be counterintuitive as the opposite is true for optical imaging.

Because range and azimuth resolution derive from different physical principles, they are not in general equal. That is why the pixels in a Single Look Complex (SLC) product are not square (see section 2.8).

## 2.4 Polarization

Like any electromagnetic radiation, radar is subject to polarization. Electromagnetic waves are *transversal* waves, i.e., the oscillation occurs in the plane perpendicular to the direction of travel. This is the opposite of *longitudinal* waves (such as sound waves). In contrast to longitudinal waves, transversal waves have a degree of freedom in terms of the angle of the field oscillation in that plane. Without going into detail, polarization is the way we are quantifying light in terms of this degree of freedom. We can describe the polarization in many different ways, but one common way is to split it into its horizontal and vertical components. All light is made up of some combination of these two components.

Depending on the platform, radar waves are emitted in either horizontal or vertical polarization, or in some cases both. Similarly, the receiving antenna may be capable of distinguishing between horizontally and vertically polarized components.

If both horizontally and vertically polarized radar is emitted and received, the system is called fully polarimetric and the resulting data are of the form

$$\Sigma = \begin{bmatrix} \sigma_{\mathrm{HH}} & \sigma_{\mathrm{HV}} \\ \sigma_{\mathrm{VH}} & \sigma_{\mathrm{VV}} \end{bmatrix} \tag{2.10}$$

. Dual polarized data, on the other hand, result from the emission of only one of the components but recording of both (as for Sentinel-1):

$$\Sigma = \begin{bmatrix} \sigma_{\mathrm{HH}} \\ \sigma_{\mathrm{HV}} \end{bmatrix} \quad \text{or} \quad \Sigma = \begin{bmatrix} \sigma_{\mathrm{VH}} \\ \sigma_{\mathrm{VV}} \end{bmatrix} \tag{2.11}$$

The backscatter components with equal emitted and received polarization are called *co-polarized* ($\sigma_{\mathrm{VV}}$ and $\sigma_{\mathrm{HH}}$), the remaining components are called *cross-polarized* ($\sigma_{\mathrm{VH}}$ and $\sigma_{\mathrm{HV}}$).

## 2.5 Scattering Mechanisms

Interpreting SAR imagery comes down to understanding the scattering mechanisms on the ground that generated the output we are observing. We may start by pointing out that most interaction of light with matter is either *transmission*, *reflection*, or *absorption*. For radar wavelengths, the atmosphere is mostly transmitting so we will ignore it for the purposes of this section. What we are really interested in is what happens to the radiation on the ground.

Surfaces differ in their proportions of reflected, absorbed, and transmitted radiation. Only the reflected portion of the signal can arrive back to the sensor (or transmitted radiation, if it is reflected at a surface below), so materials with high absorption generally appear darker. Absorption increases for example with moisture content, and is generally related to the *electric permittivity* $\varepsilon = \varepsilon_0 \varepsilon_r$ of the material ($\varepsilon_0$ is the permittivity of vacuum).

In addition to *how much* radiation is reflected, we can further differentiate between *specular* and *diffuse* reflection. A perfect specular reflector acts like a mirror and reflects light exclusively according to $\theta_i = \theta_r$ (incidence angle equal to reflected angle). A perfect diffuse reflector (also known as *Lambertian surface*) reflects incident light evenly in all directions independent of the incidence angle. In reality, no surface is perfectly specular or Lambertian, but exists somewhere on a spectrum between the two. The amount of diffuse reflection relates to the roughness of the surface as shown in fig. 2.2.

*Roughness* is always relative to the wavelength of the incident radiation. Surfaces which appear rough under visible light may appear perfectly smooth to a radar system if the features are at a scale less than the wavelength. To quantify this relationship between roughness and the wavelength we can refer to the *Rayleigh criterion*:

$$h_{\text{smooth}} = \frac{\lambda}{8 \cos \theta_i} \tag{2.12}$$

where $h_{\text{smooth}}$ is the maximum standard deviation in the height of the surface such that it still appears smooth at wavelength $\lambda$ and incidence angle $\theta_i$. The original criterion has a factor $4$ instead of $8$ in the denominator as it considers emitting sources rather than scatterers (in scattering we double the path length). Smooth surfaces such as water (which exhibit a lot of specular reflection) appear nearly black in radar imagery because for non-zero incidence angles most of the radiation

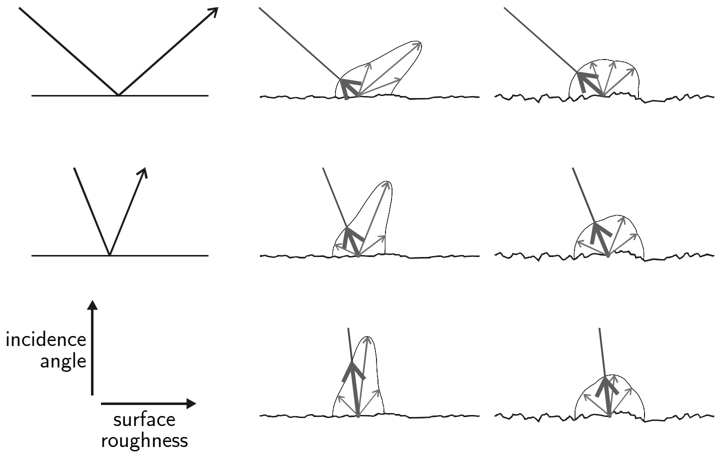


**Figure 2.2:** This figure illustrates scattering mechanisms depending on the surface roughness and incidence angle. Smooth surfaces act as specular reflectors, whereas rough surfaces are more diffuse reflectors. Graphic taken from Woodhouse (2005).

is reflected away from the sensor.

Not only the received power differs between surfaces, but also the polarization may change in the scattering process. For example, vegetation tends to depolarize radiation such that cross-polarized terms ($\sigma_{\mathrm{HV}}$ or $\sigma_{\mathrm{VH}}$) tend to be larger than for other surfaces. This is the basis of polarimetry.

## 2.6 Speckle

Speckle is a noise-like pattern that all radar imagery is subject to. It results from the fact that each resolution cell (pixel) contains many scatterers, each of which scatter a pulse with some amplitude and phase. The antenna measures only one echo, which is the superposition of all the individual backscatter signals. Some of the scattered waves interfere constructively, others destructively, generating a complex interference fringe pattern. The received power depends strongly on where in this interference pattern the antenna is located. A slight change of look angle may completely alter the received power, which is why even adjacent pixels of seemingly identical surface cover can have strong differences in their backscatter. The distribution of individual scatterers on the ground is so complex that it is impossible to predict the exact backscatter value of the pixel. Only over a larger area does the average backscatter become a value representative of that surface, which is why many speckle filters apply some form of image smoothing such as Gaussian filtering.

Yet, it is important to keep in mind that speckle is not a random effect. It is deterministic and at least in theory fully reproducible. In practice, this reproducibility is limited by the *coherence* of subsequent observations (see section 2.7).

## 2.7 Coherence

The coherence is a measure for how constant the phase difference between two waves remains over time, i.e. how well aligned their frequencies are. It is a crucial requirement for interferometry as waves only form meaningful interference patterns if they are coherent. Coherence is defined as

$$\gamma = \frac{\langle E_1 E_2^* \rangle}{\sqrt{\langle |E_1|^2 \rangle \langle |E_2|^2 \rangle}} \tag{2.13}$$

Here, $E_1$ and $E_2$ are two complex wave functions the coherence of which we are measuring. $\langle \cdot \rangle$ is the ensemble average. This may also be expressed in polar coordinates as

$$\gamma = |\gamma| e^{i\phi} \tag{2.14}$$

such that $|\gamma|$ is the coherence magnitude or correlation. In practice, the ensemble average is substituted for either a spatial or temporal average, as we are not capable of measuring multiple realizations of the same wave at the same time. If coherence is measured over time, it mostly depends on how consistent the targets are on the ground. If even a single scatterer moves between measurements, the superposition of all echos from the corresponding pixel may look completely different and the coherence drops. It follows that coherence tends to be higher for measurements with smaller time delay. Furthermore, bare ground (e.g. in deserts) exhibits higher coherence than vegetation, which is subject to continuous organic change.

While originally merely a requirement for interferometry, coherence itself may be exploited in interesting applications. In one case, researchers were able to detect vehicle paths in the desert around Nazca, Peru, by imaging the temporal coherence (Comer et al. 2017). While these paths look almost identical to the surrounding desert on the ground, the fact that gravel was displaced by even a few centimeters due to the passing vehicles caused the coherence to drop.

## 2.8 Data Acquisition and Processing

The Sentinel-1 mission consists of two satellites, Sentinel-1A and Sentinel-1B. Each satellite is in a sun-synchronous orbit at $700\,\mathrm{km}$ altitude with 175 orbits per 12-day cycle. The satellites orbit with $180^\circ$ phase difference, such that the revisit time is 6 days at the equator. Because the orbits overlap at higher latitudes, lower revisit times are achieved (about 2 days in Europe).

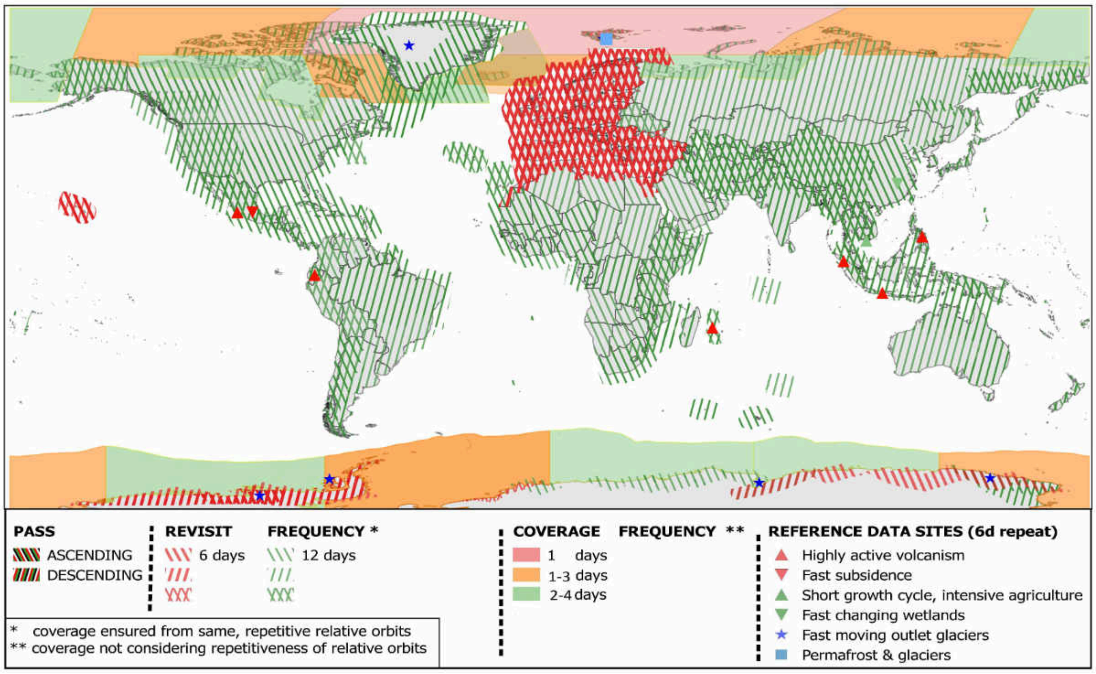


**Figure 2.3:** This figure shows the global revisit times for Sentinel-1 for ascending and descending orbits (https://sentinel.esa.int/web/sentinel/missions/sentinel-1/observation-scenario).

Sentinel-1 operates in different acquisition modes.

**Stripmap (SM)** The SM mode offers a high resolution of $5\,\mathrm{m}\times 5\,\mathrm{m}$ at a swath width of $80\,\mathrm{km}$. It is only used in exceptional cases or to image small islands.

**Interferometric Wide Swath (IW)** This is the standard mode over land. The swath width in this mode is $250\,\mathrm{km}$ with a single-look ground resolution of $5\,\mathrm{m}\times 20\,\mathrm{m}$. The image is acquired in three sub-swaths using Terrain Observation with Progressive Scans SAR (TOPSAR) (De Zan et al. 2006).

**Extra Wide Swath (EW)** This mode is very similar to the IW mode, at a higher swath width of $410\,\mathrm{km}$ and a reduced resolution of $20\,\mathrm{m} \times 40\,\mathrm{m}$. It is mostly used over maritime areas and sea ice.

**Wave mode (WV)** In WV mode, the sensor captures small patches of $20\,\mathrm{km} \times 20\,\mathrm{km}$ (called *vignettes*) at $5\,\mathrm{m} \times 5\,\mathrm{m}$ resolution. Coverage is non-continuous but occurs only every $100\,\mathrm{km}$ along the track, with alternating incidence angles in range direction.

The Sentinel-1 IW and EW products are recorded as a series of sub-swaths (three for IW and five for EW), separated in range direction and each stored as an individual image. Within each swath, the image is recorded as a series of bursts, which overlap in azimuth direction and are separated by a series of invalid pixels called *demarcation zones*. Figure 2.4 illustrates the image acquisition pattern in IW (the number of bursts in the illustration is not the actual number of bursts). An actual example showing the demarcation zones can be found in fig. 2.5.

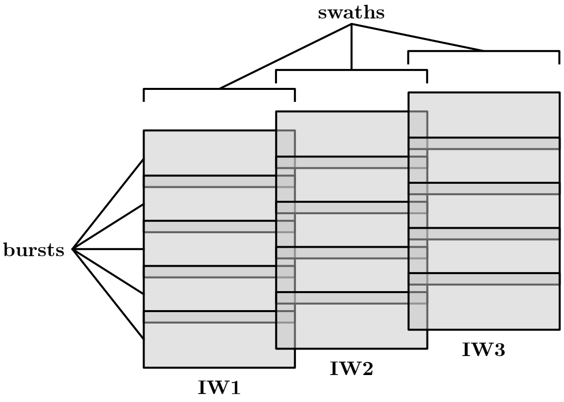


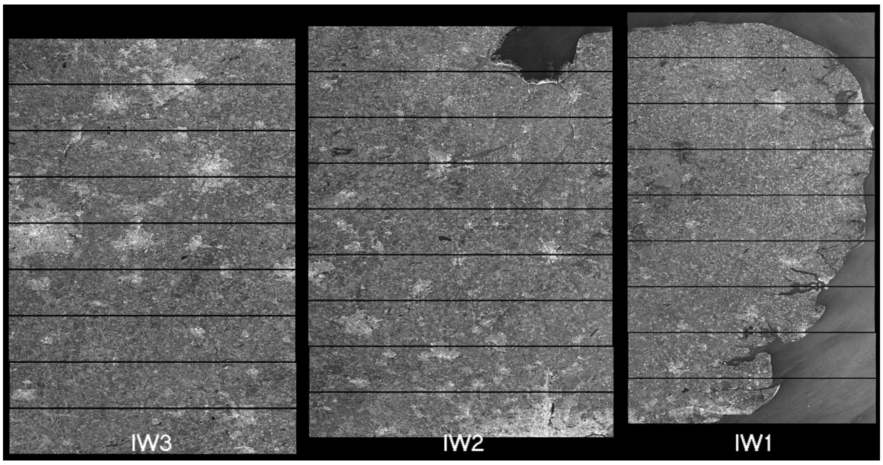
**Figure 2.4:** This graphic illustrates schematically how a single Sentinel-1 product is made up of different swaths and bursts.

The data are available at different processing levels, specifically Level-0, Level-1, and Level-2 (notation adopted from ESA). Level-0 corresponds to the raw data as obtained from the sensor and is not in general required for any application. Level-1 products have been pre-processed from the raw data and are the most relevant processing level. Level-2 is only available for ocean products.

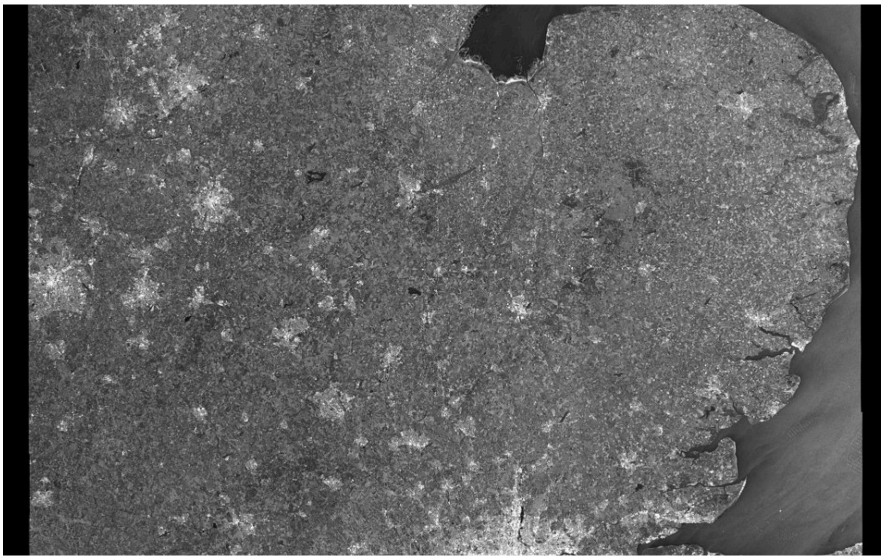
Level-1 itself is split into SLC and Ground Range Detected (GRD) products. As the name implies, SLC data are complex-valued and hence include both amplitude and phase information. GRD products do not contain any phase information and have been further processed to make them easier to analyze. On the surface they appear similar to a single-band optical image. However, as the phase information is lost, they are not suitable for interferometry and some polarimetry applications.

For this thesis, I worked with a mix of Level-1 SLC and GRD data in IW mode. All data in chapter 3 are SLC data, while in chapter 4 I mostly use GRD data. All further references to Sentinel-1 data in this chapter refer to the SLC processing level and IW acquisition mode.

Figure 2.5 demonstrates the difference between an SLC and GRD product. For comparability, the SLC product shows only the intensity; the phase information is not shown.



**(a)** IW SLC Product



**(b)** IW GRD Product

**Figure 2.5:** This figure shows the difference between an SLC product, with the swaths and bursts still visible, and a higher level GRD product. Not visible in this depiction is the fact that the SLC image contains complex valued data. Images taken from https://sentinel.esa.int/web/sentinel/user-guides/sentinel-1-sar/acquisition-modes/interferometric-wide-swath

### 2.8.1 Data access

All data originating from the Copernicus missions, including Sentinel-1, are free and open. They may be obtained from many sources through a variety of portals. The original source is the Copernicus Open Access Hub (https://scihub.copernicus.eu/dhus/) which provides both a graphical user interface and an Application Programming Interface (API).

For convenience I developed a tool to query multiple Copernicus data servers in parallel for different search constraints and perform bulk downloads of data. The tool is available from PyPI (https://pypi.org/project/esahub/).

### 2.8.2 Sentinel-1 SLC product calibration

For the work done in this research only Sentinel-1 data were used, however, the majority of the calibration steps outlined in the following are not specific to Sentinel-1 and apply to all SAR products. Sentinel-1 data over land are available as Level-0 (raw data) or Level-1 products. Level-1 encompasses both the low level SLC and a higher level GRD product. The SLC product contains both amplitude and phase information, whereas the GRD product only contains the intensity. Any application that uses Sentinel-1 data for polarimetry or interferometry thus requires the SLC product. All calibration steps can be done with the Sentinels' open source companion software SNAP.

**Apply Orbit File** Within 180 minutes of the data acquisition, ESA releases so-called Restituted Orbit (RESORB) files and 20 days later Precise Orbit Ephemerides (POEORB) files, which contain the precise location of the satellite every 10 seconds (to within $10\,\mathrm{cm}$ for the RESORB and $5\,\mathrm{cm}$ for the POEORB files) (Calero 2018, sec. 4.1 to 4.4). These orbit files are compared to the acquisition timestamp to update the orbit state vectors in the product metadata. Doing so will yield more accurate geolocation in later calibration steps.

**Radiometric Calibration** Radiometric Calibration refers to the conversion of the pixel values into calibrated radar backscatter. This step is necessary for any quantitative analysis of the data and in particular for comparability with data from other sensors.

**Deburst and Merge** These are TOPSAR specific processing steps. *Deburst* will assemble the bursts in each swath to a complete image based on the demarcation zones, and *Merge* will combine the sub-swaths.

**Speckle Filter** SAR data are subject to speckle, a noise-like pattern resulting from the superposition of the signal from multiple scatterers per resolution cell. If the radar wavelength is smaller than the resolution cell (as is usually the case), the phases of the individual scatterers appear random and contribute as either constructive or destructive interference to the measured pixel value. Speckle filters aim at reducing this effect, usually at the cost of spatial and/or temporal resolution. However, speckle is not noise in the sense of a random effect, as it is entirely deterministic and contains valuable information. For interferometric applications the speckle pattern must be retained (Lee 1983; Lee and Pottier 2009; Lee, Wen, et al. 2009).

**Terrain Flattening** Radar backscatter is affected not only by the surface type, but also by the topography, specifically its effect on the incidence angle. In order to ensure that areas of similar land cover result in similar backscatter values, Terrain Flattening compensates for backscatter variations caused by the topography such that the final result is an image as it would look if the terrain were completely flat (Small 2011). This processing step requires a Digital Elevation Model (DEM) of the region of interest.

**Terrain Correction** Whereas *Terrain Flattening* compensates for the radiometric impact, *Terrain Correction* accounts for the geometric distortions caused by the topography. The corrections use the Range Doppler orthorectification method (Small and Schubert 2008, section 4.6.2). Incidentally, in SNAP, Terrain Correction also reprojects the data into an equirectangular projection (see section 2.8.3).

**Multilooking** Multilooking is an optional processing step that reduces the image resolution by averaging over a number of adjacent pixels. If high-resolution data are not of interest, this is usually done to either reduce speckle or achieve a nominal image pixel size. However, both effects may be achieved by a combination of a speckle filter and terrain correction, such that a multilooking step is not in general necessary and does not justify the loss of resolution.

### 2.8.3 Geocoding

Geospatial data are usually provided with geocoding information in one of two ways. One option is a list of Ground Control Points (GCPs), where for a small subset of the pixels the latitude, longitude, elevation, etc. are given. If these points are organized in a grid this is sometimes referred to as a *tie point grid*. It is then up to the data processing application to use this information, e.g. by interpolating the coordinates for the remaining pixels.

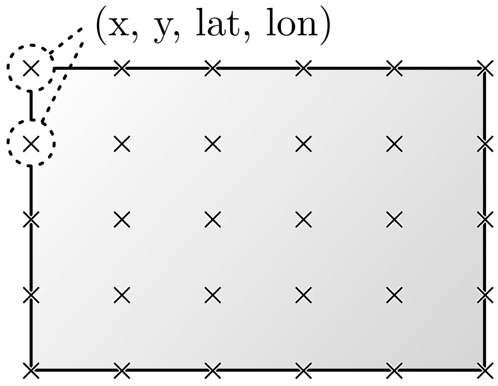


**Figure 2.6:** An illustration of GCPs.

A second way in which coordinates may be encoded is with a so-called *affine transformation* $A$, which is typically given as a set of 6 values corresponding to

$$\begin{bmatrix}\theta_{lat}\\ \theta_{lon}\end{bmatrix} = A \cdot \begin{bmatrix} y \\ x \\ 1 \end{bmatrix}, \quad A \in \mathbb{R}^{2\times 3} \tag{2.15}$$

where $\theta_{lat}$ and $\theta_{lon}$ are the latitude and longitude, and $y$ and $x$ are the image coordinates, i.e. pixel row and column.

The identity transform is

$$A_I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}. \tag{2.16}$$

In general, $A_{11}$ and $A_{22}$ indicate the scaling or stretching, $A_{12}$ and $A_{21}$ are non-zero if the transform contains elements of shear or rotation, and $A_{13}$ and $A_{23}$ encode the offset.

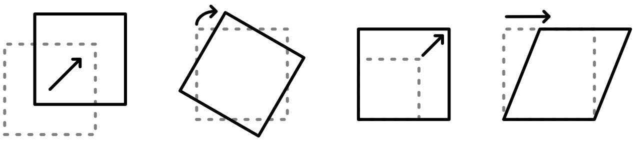


**Figure 2.7:** An affine transformation may contain the following elements (from left to right): translation, rotation, scaling, shear.

Six parameters are thus required to define the transformation: $A_{11}$, $A_{12}$, $A_{13}$, $A_{21}$, $A_{22}$, $A_{23}$. This is the natural order in which e.g. the affine Python library requires the coefficients. However, other standards for the order of the coefficients exist:

| | | |
|---|---|---|
| affine Python library | $A_{11}, A_{12}, A_{13}, A_{21}, A_{22}, A_{23}$ | |
| ESRI World File | $A_{11}, A_{21}, A_{12}, A_{22}, A_{13}, A_{23}$ | |
| GDAL | $A_{13}, A_{11}, A_{12}, A_{23}, A_{21}, A_{22}$ | (corresponding to reordering the image coordinate vector as $[\,1\ y\ x\,]^T$) |

**Table 2.1:** Ordering standards for the coefficients of an affine transformation.

As the affine transformation only supports linear transformations – and the relation between a recorded image and geodetic coordinates is rarely linear – the use of such a transform usually indicates that the image has already been reprojected (or covers a very small area such that a linear fit may be a reasonable approximation).

In order to get the data into an analysis ready format, it is usually most convenient to resample the data on a grid where longitude and latitude correspond to the $x$ and $y$-axis, respectively. This projection is known as the *equirectangular* projection and is associated with the code EPSG:4326 (European Petroleum Survey Group, EPSG) or WGS84 (World Geodetic System, WGS). It is especially easy to deal with as the coordinate axes directly correspond to the latitude and longitude dimensions.

However, the main drawback of this projection is the distortion of the pixels towards the poles. At the equator, one degree latitude and one degree longitude correspond to approximately the same distance on the ground ($111$ km). At increasing latitude, the circles representing $360°$ get smaller. The ratio between the distance covered by a degree latitude and a degree longitude is given by

$$\frac{d_{lon}}{d_{lat}} = \cos(\theta_{lat}) \tag{2.17}$$

such that in Edinburgh ($\theta_{lat} \approx 56°$), one degree longitude corresponds to only about $62\,$km. Therefore, if we choose square pixels in lat-lon space, at $60°$N there will be twice as many pixels per unit length on the ground in the East-West direction as for the North-South direction. If the study area is restricted to a small geographic footprint, this does not typically pose a problem because the resolution of the reprojected image can be accurately matched to the native resolution of the unprojected image by choosing appropriate scales for latitude and longitude independently (alternatively, one can choose a different projection that is tailored to the geographic area of interest).

For global datasets, however, a common resolution in lat-lon space is required when using this projection, resulting in vastly different resolutions on the ground. Assuming the native resolution of the imagery is the same near the equator and near the poles, the reprojected (resampled) dataset will be either oversampled near the poles or undersampled near the equator. As a consequence, other projections may be more appropriate depending on the application, especially equal-area projections such as the Albers or Mollweide projections.

### Data reprojection

The terms *reprojection*, *warping*, and *resampling* are often used interchangeably so it is a good idea to briefly define them more clearly. For image data, resampling refers to the process of computing data values at different coordinates, for example through aggregation or interpolation. This process may change the image size, as in up- or downsampling. Reprojection refers to resampling in a specific coordinate system (projection). While there are many different coordinate systems commonly used in geospatial data, we will largely work with EPSG:4326 as discussed previously. Although image warping is more generally referring to any form of image deformation or distortion, it is used here synonymously with reprojection, mostly due to its use in the tool *gdalwarp* (GDAL Development Team 2018).

Resampling data to a new coordinate grid requires two steps:

1. **Coordinate transform**. For each pixel in the new grid, the corresponding image coordinates of the original image need to be determined.
2. **Data interpolation**. The pixel values at the required image coordinates are then evaluated. As these are not in general integer coordinates, interpolation is necessary.

The coordinate transform is straightforward in the case where an affine transformation is given, or where the exact latitude and longitude are known for each pixel. If only a list of GCPs is known, this step also involves an additional interpolation step to find the coordinates at the remaining pixels.

First, the GCPs are interpolated to give a function that maps from lat-lon space to the image coordinates $x, y$:

$$f : \mathrm{lat}, \mathrm{lon} \mapsto x, y$$

The image has a number of data values attached to each pixel, which we will denote as $\boldsymbol{\phi}(x, y)$. This function is known at discrete points corresponding to the image pixels and can be interpolated for the coordinates in between.

Next, an output grid needs to be defined:

$$g : x', y' \mapsto \mathrm{lat}, \mathrm{lon} \quad \text{for} \quad x' \in [0, x'_{max}], y' \in [0, y'_{max}]$$

For our equirectangular projection, this is a rectangular grid in lat-lon space such that $x' \propto \mathrm{lon}$ and $y' \propto \mathrm{lat}$.

Finally, $f$ and $g$ can be combined to get a coordinate transform that maps from the new to the old image coordinates.

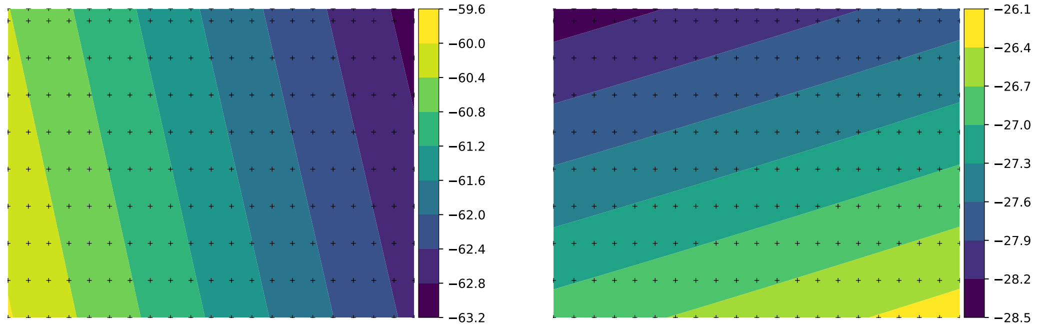
$$f \circ g : x', y' \mapsto x, y$$

Using this coordinate transform, the data values can be mapped to the output grid:

$$\boldsymbol{\phi}'(x', y') = \boldsymbol{\phi}((f \circ g)(x', y'))$$

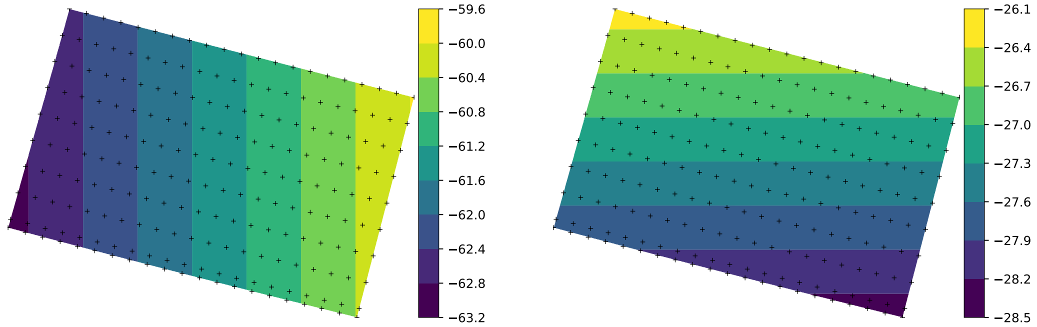
Figure 2.8 demonstrates this process by showing the interpolated longitude and latitude values of the original image (figs. 2.8a and 2.8b) as well as the warped image (figs. 2.8c and 2.8d). The positions of the GCPs are marked in black in each image. The coordinates shown here correspond to the (arbitrary) product

`S1B_IW_GRDH_1SDV_20170107T092836_20170107T092901_003740_0066CF_CB57`



**(a)** Original longitude (interpolated from GCPs)

**(b)** Original latitude (interpolated from GCPs)



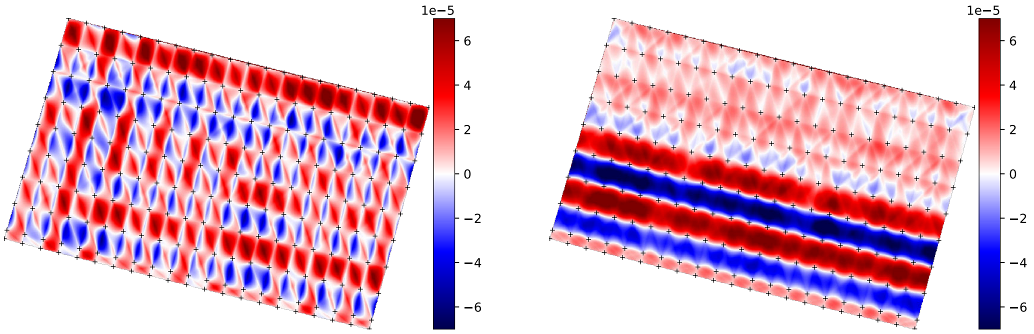
**(c)** Warped longitude

**(d)** Warped latitude

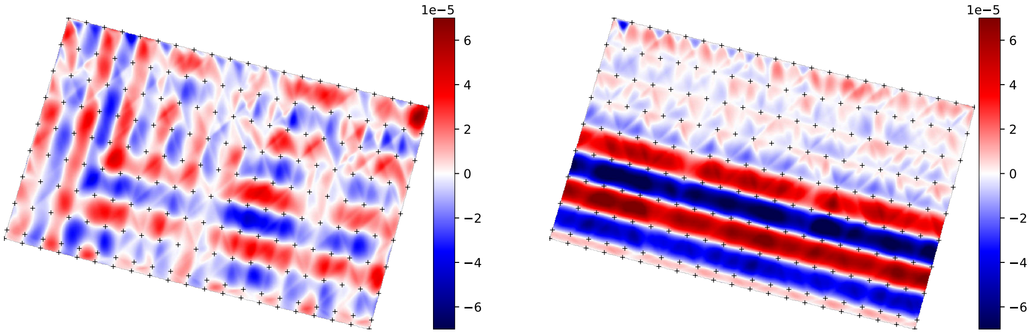
**Figure 2.8:** This figure shows the distribution of latitude and longitude in the original image as well as in the warped/reprojected image. GCPs are marked as black crosses.

While the latitude and longitude appear roughly linear in the original image coordinates, this is not in general the case. Interpolation based resampling yields images (c) and (d), where longitude and latitude correspond to the $x$ and $y$ axis, respectively. Resampling of actual data is done in exactly the same way.

Because in reality most raw data are geocoded with GCPs, resampling to the equirectangular projection thus requires two interpolation steps. It is worthwhile at this point to assess the resulting interpolation error. It is impossible to assess the error in terms of how well the data are interpolated because we have no information about the data beyond the pixel values. However, we can estimate the error resulting from the reprojection by looking at the accuracy of the spatial resampling itself. We can get an idea of this error by reprojecting the interpolated latitude and longitude coordinates as shown in fig. 2.8. Ideally, the coordinates should line up with the predefined output grid. However, as there are several interpolation steps involved, a deviation of the warped coordinates from the reference grid can be observed. The differences are shown in fig. 2.9 for linear and cubic interpolation.



**(a)** Interpolation error in longitude, linear interpolation **(b)** Interpolation error in latitude, linear interpolation



**(c)** Interpolation error in longitude, cubic interpolation **(d)** Interpolation error in latitude, cubic interpolation

**Figure 2.9:** This figure shows the interpolation errors in latitude and longitude derived from the difference between the warped coordinates and the reference grid.

The plot shows that the interpolation error with this definition is in the order of $10^{-5}$ degrees latitude and longitude, which at the equator is in the order of meters. At $10\,\mathrm{m}$ resolution this would mean that the uncertainty in the geolocation is a little less than one pixel. For many applications this error is acceptable and probably not even noticeable. However, it is important to keep in mind that this error exists and is potentially multiplied the more interpolation/resampling steps an image undergoes. Furthermore, this assessment does not take into account any error that the GCPs themselves may carry.

The main scenario where these errors do play a significant role is in interferometry. Interferometry relies on very high coherence between pixels over time, i.e. consistent phase information. As the phase may change completely between adjacent pixels, even sub-pixel errors cause interferometry to fail. This creates a need for highly accurate image alignment algorithms (in remote sensing this is called *coreg-*

*istration*). These algorithms align images based on e.g. feature matching, DEM alignment or spectral techniques such as *enhanced spectral diversity* (Scheiber et al. 2000) to achieve coregistration with sufficient sub-pixel accuracies.

Coregistration algorithms do not eliminate the geocoding error. Rather, they align the geocoding of a stack of images with respect to a master image, such that the same pixel in every image refers to the same area on the ground.

## 2.9 Summary

In this chapter, I have presented an overview of the theoretical concepts behind SAR imaging by differentiating it from optical imaging in particular. The key differences between optical and radar imaging can be summarized as follows:

1. Radar is an active system, whereas optical sensors are passive.
2. The wavelength of a radar system is much longer at $1$ cm to $1$ m (compared to $10$ nm to $1$ mm for optical sensors).
3. In a radar system, the reflected intensity is measured as a function of range rather than look angle.

Any differences in the specific application of either radar or optical sensors may be attributed to one or multiple of these key elements. For example, the longer wavelength of a radar instrument is similar in order of magnitude to the size of objects we are trying to measure. For this reason, radar can provide more direct information about the three-dimensional structure of objects on the ground, which is useful for measuring vegetation related quantities such as forest height and biomass. The longer wavelength has the additional advantage that it does not interact with much smaller objects, such as the liquid droplets that form clouds, which is why radar is unaffected by clouds. Because radar instruments are active systems, they also provide their own source of illumination. They are therefore not only independent of atmospheric conditions, but also of time of day and general lighting conditions, making them ideal for change detection.

# Chapter 3

# Forest/Non-Forest Separability

## 3.1 Introduction

Detecting changes in the forest cover using SAR data requires a statistically significant difference in the backscatter signature of forested and non-forested areas. This chapter assesses empirically how well forested areas can be distinguished. At its core this is a classification problem and as such we begin by reviewing a number of papers that have attempted to classify forest and non-forest based on SAR data.

Shimada, Itoh, et al. (2014) present a new forest/non-forest map with global coverage derived from ALOS PALSAR (L-band SAR). They note that most existing forest cover maps have either been restricted to less-than-global coverage due to focus on specific forest types or geographic regions or been generated for a single time. Furthermore, all previous global forest maps have been derived from optical imagery (e.g. Landsat) and as such are subject to cloud cover limitations. This requires a mosaic over multiple time steps for a cloud-free composite and hence limits the temporal repeatability. Two notable exceptions of repeated global forest maps are the maps by Hansen, Potapov, et al. (2013) and the FAO (Lindquist et al. 2012). The authors further explain the particular aptitude and limitations of L-band SAR for forest mapping. While non-forest areas typically yield lower backscatter than forest, deforestation detection is hindered by an initial increase in backscatter following forest clearance because of remaining debris as well as rainfall. The detection potential was thus found to be greatest once all debris was removed and the area had been converted into agriculture.

The final data product used for classification was a $25\,\mathrm{m}$ global mosaic with 16 looks, orthorectified using the SRTM DEM (Shimada and Ohtaki 2010) and expressed as gamma-nought ($\gamma_0$), i.e. the incidence angle normalized radar cross section:

$$\gamma_0 = \frac{\sigma_0}{\cos\theta} \tag{3.1}$$

Prior to classification the data were smoothed using a $5 \times 5$ median filter.

Training and validation data were derived from Google Earth imagery, the DCP, and the FRA (Lindquist et al. 2012). The authors used a forest definition of more than 10% woody vegetation cover throughout. Analysis of forest/non-forest histograms showed that forests followed a normal distribution in $\gamma_0$ with the co-polarized backscatter always being larger than the cross-polarized component. They derived the cross-over thresholds to be $-7.5\,\mathrm{dB}$ and $-14.0\,\mathrm{dB}$ for HH and HV, respectively. The backscatter over forest tended to follow a normal distribution with $-6.89\,\mathrm{dB}$ (standard deviation $2.13\,\mathrm{dB}$) for HH and $-13.07\,\mathrm{dB}$ (standard deviation $2.04\,\mathrm{dB}$) for HV. For non-forest areas, the backscatter was lower and less consistent: $(-10.86 \pm 4.78)\,\mathrm{dB}$ and $(-18.49 \pm 3.84)\,\mathrm{dB}$ for HH and HV, respectively. An overall decrease in $\gamma_0$ was noted for the observation period from 2007 to 2010, which the authors attribute to decreases in forest cover and biomass.

The authors used a very simple manually crafted decision tree as classifier, relying on region-specific backscatter thresholds extracted from forest/non-forest histograms. They further decided to only use HV polarization for classification.

The total forest cover obtained this way was about 5% less than the FRA and Hansen results. Validation showed an accuracy of 91.25% compared with Google Earth imagery, 94.81% compared with the FRA, and 84.86% compared with the DCP. Discrepancies were attributed to different image acquisition dates, cloud cover, confusion of woody vegetation with herbaceous vegetation in maps derived from optical data, as well as too coarse resolution to detect forest with low cover.

Quegan, Toan, et al. (2000) assess the potential of time series of C-band SAR images for forest mapping. They note that C-band SAR has been less popular for forestry applications because backscatter at this relatively short wavelength saturates at biomass densities of $30$ to $50\,\mathrm{t\,ha^{-1}}$, whereas for P-band SAR saturation only happens around $150$ to $200\,\mathrm{t\,ha^{-1}}$, making it capable of assessing variations in a larger range of forest types. Different saturation thresholds exist for different species due to differences in their dielectric properties, as well as leaf size and orientation. This makes C-band especially ill-suited for detecting forest degradation, i.e. a reduction of biomass while maintaining forest status, as in many parts of the world forests can exceed a biomass density of $500\,\mathrm{t\,ha^{-1}}$.

One incentive for using multi-temporal SAR data is that at individual time steps the backscatter can often look identical for forest and other vegetation types. However, because of seasonal effects certain times of year may provide higher contrast. There is a relationship between forest stand age and backscatter, as with growing age the attenuation of the soil contribution increases, resulting in a reduction of the overall backscatter. Following the initial decrease in backscatter for a growing forest there is a slower increase leading up to the eventual saturation at the biomass limit.

Because denser forest exposes less of the soil, the backscatter is less dependent on soil moisture and roughness and hence exhibits less temporal variability than sparser vegetation or bare soil. Other stable land cover includes grassland and urban areas, but these are stable at lower and higher backscatter values, respectively.

Since non-forest areas exhibit greatly varying backscatter values over the year, they occasionally look similar to forest. For these reasons, the authors propose temporal backscatter stability as the most promising feature for forest mapping using C-band SAR data. The processing chain consists of typical preprocessing, speckle reduction, and classification steps. The speckle reduction is done with consecutive temporal and spatial filters to achieve a high Estimated Number of Looks (ENL) of greater than 100 which was required to avoid excessive false positive rates at change amplitudes of $5\,\mathrm{dB}$.

The following temporal filtering was applied:

$$J_k(x,y) = \frac{\langle I_k \rangle}{N} \sum_{i=1}^{N} \frac{I_i(x,y)}{\langle I_i \rangle} \qquad 1 \leq k \leq N \tag{3.2}$$

with the assumption that temporal correlation is near zero. Subsequently, a spatial filtering with an $11 \times 11$ window was applied. Using the pre-processed and filtered images, classification was then done based on the *mean annual variation* (mva):

$$\mathrm{mva} = 10\log\left[\frac{2}{N(N-1)}\sum_{i=1}^{N-1}\sum_{j>i} R_{ij}\right] \tag{3.3}$$

where $R_{ij} = \max(I_i/I_j, I_j/I_i)$.

Two types of land cover were identified that would yield classification errors for a simple thresholding of the mva: non-forest areas of low backscatter variability (e.g. built up areas or grassland), and forest areas of high backscatter variability. The classification accuracies for three test sites in the UK (11 images), Finland (15 images), and Poland (2 images) were $90.7\%$, $94.1\%$, and $77.4\%$, respectively.

The authors conclude that temporal backscatter stability is in fact the best predictor for forest cover derived from C-band SAR for forest of more than $30\,\mathrm{t\,ha^{-1}}$.

McRoberts et al. (2007) elaborate on the potential of remote sensing data for national forest inventories (NFIs). They emphasize the low cost of measuring areal quantities such as forest area or volume compared to conventional field sampling, placing a particular focus on k-Nearest-Neighbors (kNN) methods. However, they note limitations in quantifying the uncertainty of forest inventories.

## 3.2 Methods and Data

### 3.2.1 Study sites

We have selected six study sites across a broad range of forest biomes: Alaska, Florida, Finland, UK, Indonesia, and Colombia. The locations of the study sites are shown in fig. 3.1 and in more detail in fig. 3.2.

The UK site is an area in the *Kielder Forest Park* for the year 2015 and is largely covered by coniferous woodland, grassland, and bog. Reference data for this site are taken from the 2015 land cover map by the UK Centre for Ecology and Hydrology (https://www.ceh.ac.uk/services/land-cover-map-2015).

The site in Florida is part of the *Ocala National Forest* for the year 2017. This site exhibits diverse land cover including (in order of prevalence) sand pine scrub,

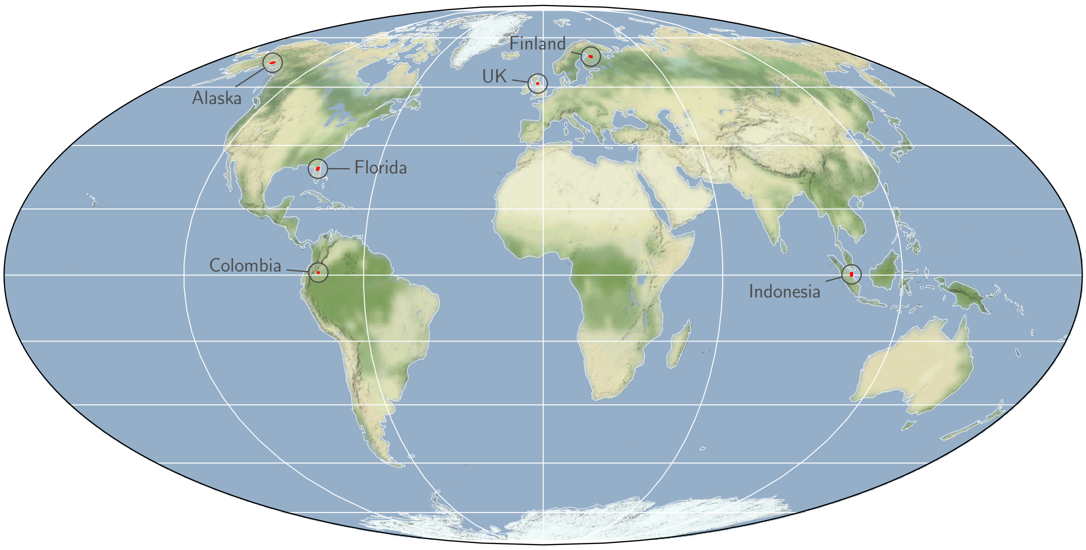


**Figure 3.1:** This figure shows all six areas that were chosen for the separability study on a single map. The exact locations are marked with red polygons, with annotations to help locate them visually. The map was created using Cartopy (Met Office 2010 - 2020) with map tiles by Stamen Design, under CC BY 3.0.

coniferous plantations, lacustrine, pasture, sandhills, marshes and swamps, as well as some urban areas (including transport infrastructure) and water bodies. The land cover reference is taken from the Florida Cooperative Land Cover Map 2017 (https://myfwc.com/research/gis/applications/articles/cooperative-land-cover/).

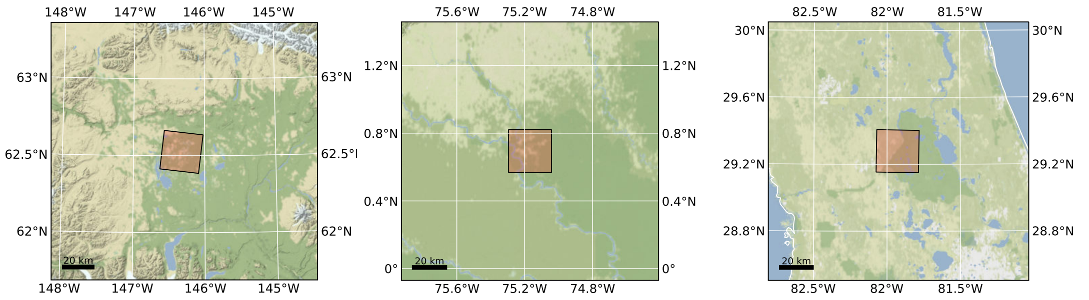
The Alaskan site is located around *Lake Louise* for the year 2017 and is mostly covered by spruce, as well as some (low) shrub and freshwater bodies. The corresponding reference data set is the Alaska Vegetation and Wetland Composite 2017 (https://accscatalog.uaa.alaska.edu/dataset/alaska-vegetation-and-wetland-composite).

The Indonesian site is the *Indragiri Hilir Regency* for the year 2017. The most common land cover classes in this area are shrub-mixed dryland farm, swamp shrub and swamp forest, bare land, various plantations, and mangrove forest. A 2017 land cover map by the Indonesian Ministry of Environment and Forestry is used for reference (http://data.globalforestwatch.org/datasets/land-cover-indonesia).

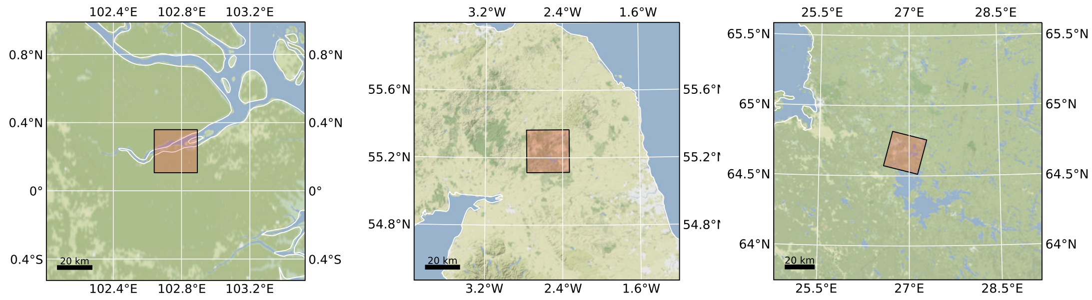
The Colombian site is located near *Solano, Caquetá* for the year 2018 and is covered by tall dense forest, rivers, and pasture. The reference data for this site were obtained from a national forest inventory data set consolidated by the Colombian Institute of Hydrology, Meteorology and Environmental Studies (IDEAM) (http:

//smbyc.ideam.gov.co/).

For Finland, a site just north of *Lake Oulujärvi* was selected for the year 2018. The majority of this site is covered by coniferous forest, with the remaining area made up of water bodies, transitional woodland-shrub, mixed forest, and peat bogs. Reference data were derived from the CORINE Land Cover map 2018 (https://land.copernicus.eu/pan-european/corine-land-cover/clc2018).



**(a)** Alaska **(b)** Colombia **(c)** Florida



**(d)** Indonesia **(e)** UK **(f)** Finland

**Figure 3.2:** This figure shows the location of the six study areas marked in fig. 3.1 in more detail.

### 3.2.2 Reference data

The reference data for each study site were constructed from the GlobalForestWatch data set (Hansen, Potapov, et al. 2013) as well as a country-specific land cover map for the corresponding year. It is assumed that both data sets contain inaccuracies for different reasons: The Hansen data set is derived from a Machine Learning approach and hence may not take into account local expert knowledge about differences between biomes. The country specific landcover maps, on the other hand, are released less frequently and only provide a snapshot at a certain time. They hence do not capture land cover change during the study period. For these reasons, only pixels where both data sets agree were used as reference throughout the study. Figure 3.3 shows the Hansen maps as well as country specific reference data for each of the sites. The third row indicates the areas of agreement between the two. The percentages of agreement are listed in table 3.1. The last two rows show Sentinel-1 images for January and December of the respective year. Note that the Colombian reference data set is the only one that provides a deforestation class, which is why the image labeled "*country reference*" shows a few red patches for Colombia.

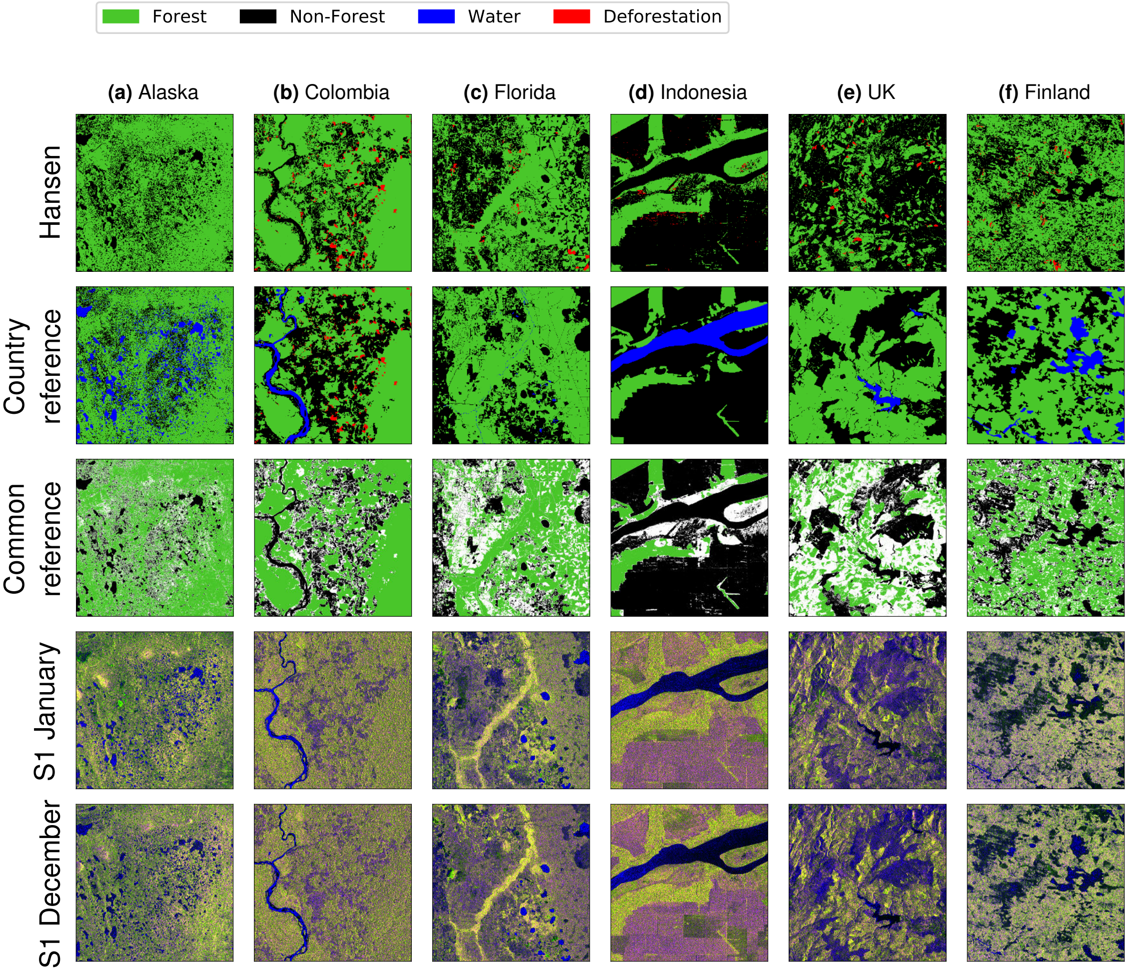


**Figure 3.3:** This figure shows the reference data for each of the six study sites, as well as Sentinel-1 images of the start and end of the respective year. Forested areas are shown in green, non-forest in black, deforestation in red, and water in blue. Areas of disagreement between the two references are left white in the third row. The Sentinel-1 images show the intensity in VV polarization in the red channel, VH in the green channel, and their ratio in the blue channel.

**Table 3.1:** This table shows the agreement between the Hansen dataset and the country specific land cover dataset for each study site.

| | **Alaska** | **Colombia** | **Florida** | **Indonesia** | **UK** | **Finland** |
|---|---|---|---|---|---|---|
| agreement | 77.1% | 83.3% | 64.7% | 86.5% | 58.8% | 76.9% |

### 3.2.3 Data processing

All data analyzed in this study are Sentinel-1 SLC products. The data were preprocessed with ESA SNAP by applying the following processing steps.

**Apply Orbit File** All products were geocoded using Precise Orbit Determination (POD) files with polynomial degree 3.

**Radiometric Calibration** The products were radiometrically calibrated to $\sigma_0$, preserving complex output (Rosich et al. 2004).

**Debursting** In this step, the individual bursts within each subswath were assembled.

**Polarimetric Speckle Filter** For speckle reduction, a Refined Lee Filter with window size $7 \times 7$ was applied (Lee 1981; Lee and Pottier 2009).

**Terrain Correction** The products were processed to $14.07$ m resolution using Range Doppler Terrain Correction with bilinear resampling (Small and Schubert 2008). This is the native azimuth resolution of Sentinel-1 in IW mode and the resulting resolution of square pixels without multilooking. The terrain correction was done using the SRTM 1 arc-second DEM. As the Alaskan site is north of $60°$N it is not covered by the SRTM and a separate DEM was used, in this case the *Alaska 2 Arc-second Digital Elevation Model* provided by the United States Geological Survey (USGS). The products were projected to the following map projections: Alaska: *EPSG:3338*, Colombia: *EPSG:3116*, Finland: *EPSG:3035*, Florida: *EPSG:3086*, Indonesia: *EPSG:4326*, UK: *EPSG:27700*

Because we are working with time series data there is scope for temporal smoothing. Multilooking was therefore not deemed necessary, allowing us to preserve spatial resolution. While the unprocessed SLC data can be represented by a complex scattering vector $\boldsymbol{s}$ containing the amplitude and phase of the co-polarized (VV) and cross-polarized (VH) backscatter

$$\boldsymbol{s} = \begin{bmatrix} S_{VV} \\ S_{VH} \end{bmatrix} = \begin{bmatrix} A_{VV}e^{i\varphi_{VV}} \\ A_{VH}e^{i\varphi_{VH}} \end{bmatrix}, \tag{3.4}$$

we used the C2 representation of the data, which is the standard output of the preprocessing chain and the basis for any polarimetric applications:

$$C = \boldsymbol{s}\boldsymbol{s}^H = \begin{bmatrix} A_{VV}^2 & A_{VV}A_{VH}e^{i\Delta\varphi} \\ A_{VV}A_{VH}e^{-i\Delta\varphi} & A_{VH}^2 \end{bmatrix} \tag{3.5}$$

where $(\cdot)^H$ denotes the Hermitian transpose: $(A^H)_{ij} = (A_{ji})^*$, and $\Delta\varphi = \varphi_{VV} - \varphi_{VH}$.

As a consequence of this representation, the absolute (random) phase is no longer available. As the dependency on the absolute phase has been removed, we cannot reconstruct the scattering vector from the covariance matrix. This leaves three independent variables: $C_{VV}$ (the intensity in VV), $C_{VH}$ (the intensity in VH), as well as $\Delta\varphi$ (the VV-VH phase difference). In the case of fully polarimetric data, the covariance matrix would contain additional terms that include the co-polarized phase difference (VV-HH). While the co-polarized phase difference is related to the number of bounces, and as such is often useful for land cover classification, the cross-polarized phase difference is generally regarded as noise (Lee and Pottier 2009). For the dual-polarized Sentinel-1 data, the HH component and thus the co-polarized phase difference is not available. In this study we compare the ability of these three available variables, individually and together, to distinguish between forest and non-forest. As part of this analysis we test the assertion that the cross-polarized phase difference contains no useful information.

Due to the way the satellites orbit the Earth, roughly half of the data are acquired while the satellite is moving from south to north (this is called *ascending* orbit) and the other half when it is moving from north to south (*descending* orbit). This results in different look angles which have a particular impact on the appearance of terrain with topography. This effect can be partially accounted for using a processing step called *terrain flattening*. However, we do not apply this step as it does not preserve the phase information. Instead, throughout this study, we treat the measurements in both orbits as separate variables. Aggregating over them would artificially increase the variability over pixels affected by topography which would reduce the separation between forest and non-forest distributions.

### 3.2.4 Distribution separability

To assess the separability of forest and non-forest pixels with Sentinel-1 data, we first looked at the distribution of backscatter values within each class. In order to quantify the aptitude of each feature for separating between forest and non-forest, we can compute the Overlap Coefficient (OVL) of the forest and non-forest distributions for each variable. Given two probability density functions (PDFs) $f(x)$ and $g(x)$, the OVL is defined as

$$\text{OVL}(f, g) = \int_{\mathbb{R}^n} \min\left(f(x), g(x)\right) dx \tag{3.6}$$

which can be approximated for the discrete case as

$$\text{OVL}(f, g) = \sum_{i=0}^{N} \min\left(f_i, g_i\right) \tag{3.7}$$

where $f_i$ and $g_i$ are the values of $f(x)$ and $g(x)$ known only at discrete positions $x_i$ for $i = 0, \ldots, N$ normalized such that $\sum_i f_i = \sum_i g_i = 1$. An OVL of 1 indicates equality of the distribution, whereas an OVL of 0 means zero overlap and therefore perfect separability.

To see how the theoretical separability holds up in practice, a number of classifiers were trained on the labeled data and their results and scores are compared. The training data comprised a random selection of $10\,000$ data points per class to discriminate between forest and non-forest. The points were selected at random out of those pixels where both reference data sets are in agreement. All validation was equally performed with respect to the areas of agreement between both references only; areas of disagreement were excluded for both training and validation. There are three available variables: $C_{VV}$, $C_{VH}$, and $\Delta\varphi$. From these variables five different feature sets were generated:

1. Each variable measured once in ascending orbit and once in descending orbit (six features in total).
2. Each variable measured at every time step (3 × [number of time steps] features in total).
3. The mean and standard deviation of each variable, split between ascending and descending orbit
    (a) for all three variables (six features in total).

(b) for $C_{VV}$ and $C_{VH}$ only (four features in total).

(c) for $\Delta\varphi$ only (two features in total).

The classifiers tested were k-Nearest-Neighbors (kNN), Random Forest (RF) (Breiman 2001), Quadratic Discriminant Analysis (QDA), and Support Vector Machine (SVM) (Cortes et al. 1995) to cover the most common classifiers in the remote sensing community. QDA is not generally used on a pixel-by-pixel basis as a classifier, but was added as a simple baseline model for comparison. Prior to classification, the features were standardized by subtracting the mean and scaling to unit variance. The scaling is of particular importance for the kNN and to some extent the SVM classifier.

Each classification outcome was rated using the balanced accuracy score as well as Cohen's Kappa score. The balanced accuracy score is an adaptation of the standard accuracy score which accounts for class imbalances, i.e., the case where the available classes have different numbers of members (Brodersen et al. 2010). It is defined as

$$\text{BACC} = \frac{1}{2}\left(\frac{\text{TP}}{\text{TP+FP}} + \frac{\text{TN}}{\text{TN+FN}}\right) \tag{3.8}$$

with TP, FP, TN, and FN being the number of *true positive*, *false positive*, *true negative*, and *false negative* observations, respectively.

Cohen's Kappa is a measure of how much better a classifier performs compared with random guessing (Cohen 1960). It is defined as

$$\kappa = \frac{p_0 - p_e}{1 - p_e} \tag{3.9}$$

where $p_0$ is the standard accuracy and $p_e$ is the probability of agreement by chance which can be computed as

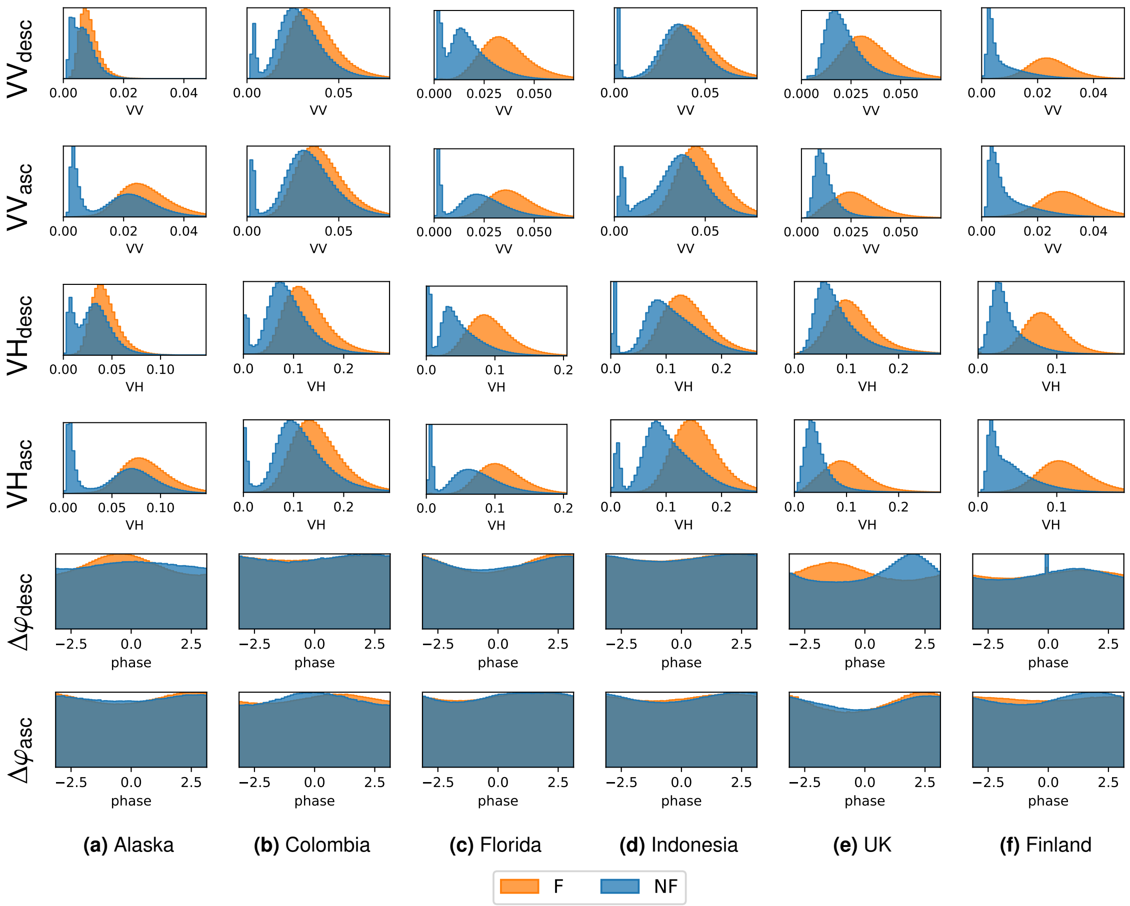
$$p_e = \frac{1}{N^2}\sum_k n_k n'_k \tag{3.10}$$

for $N$ observations and $k$ classes, with $n_k$ being the true number of members of class $k$ and $n'_k$ the predicted number of members of class $k$. A score of $0$ indicates no improvement over random guessing. Just like the standard and balanced accuracy scores, the maximum score of $1$ is reserved for perfect agreement. The Kappa score allows for negative scores, indicating performance worse than random guessing.

## 3.3 Results

### 3.3.1 Data distribution

Figure 3.4 shows the distribution of backscatter intensity in VV and VH as well as the phase difference $\Delta\varphi$ split between ascending and descending orbits for each of the study sites, per class (forest and non-forest). Similarly, fig. 3.5 shows the distribution of annual mean and standard deviation of these values. We chose to show the distributions of temporal mean and standard deviation rather than the unaggregated data to yield more pronounced differences between sites and land cover classes as well as mitigate the underlying variability in the backscatter values either due to noise, seasonality, or other effects.



**(a)** Alaska **(b)** Colombia **(c)** Florida **(d)** Indonesia **(e)** UK **(f)** Finland

**Figure 3.4:** This figure shows the distribution of the co-polarized and cross-polarized backscatter (VV and VH) as well as the phase difference ($\Delta\varphi$) per land cover class (forest, non-forest) according to the reference data, for a single image. The distributions are shown separately for the descending and ascending orbits, as denoted by the subscripts “asc” and “desc”. Forest distributions are shwon in orange, non-forest distributions in blue.

As a general observation, both VV and VH backscatter are larger for forested areas than for non-forested areas, on average. However, the individual distributions appear very different across different study sites. For Alaska, the distributions are nearly identical, whereas for Finland and the UK, the distributions separate very well. An interesting observation can be made about the phase difference. Overall, it seems that there is essentially no information contained in the phase difference when it comes to distinguishing forest from non-forest. The exception is the UK study site in descending orbit, and to some extent Finland, where there is a clear shift between the distributions for the two classes. Less apparent differences are also visible for Alaska and Colombia. The distributions of the temporal mean show similar results compared to the distribution of the raw data. The differences appear in general more pronounced, indicating a mitigation of the underlying variability in the backscatter values either due to noise, seasonality, or other effects. Contrary to what has been found by Quegan, Toan, et al. (2000), the histograms of the temporal standard deviation of VV and VH backscatter seem to suggest that forested areas have a larger rather than smaller temporal variability compared to non-forested areas. This is in particular true for regions with pronounced seasonality such as the UK, Alaska, and also Florida, but also holds true to a lesser extent for Colombia and Indonesia. The standard deviation of the phase difference shows no difference between the two classes.

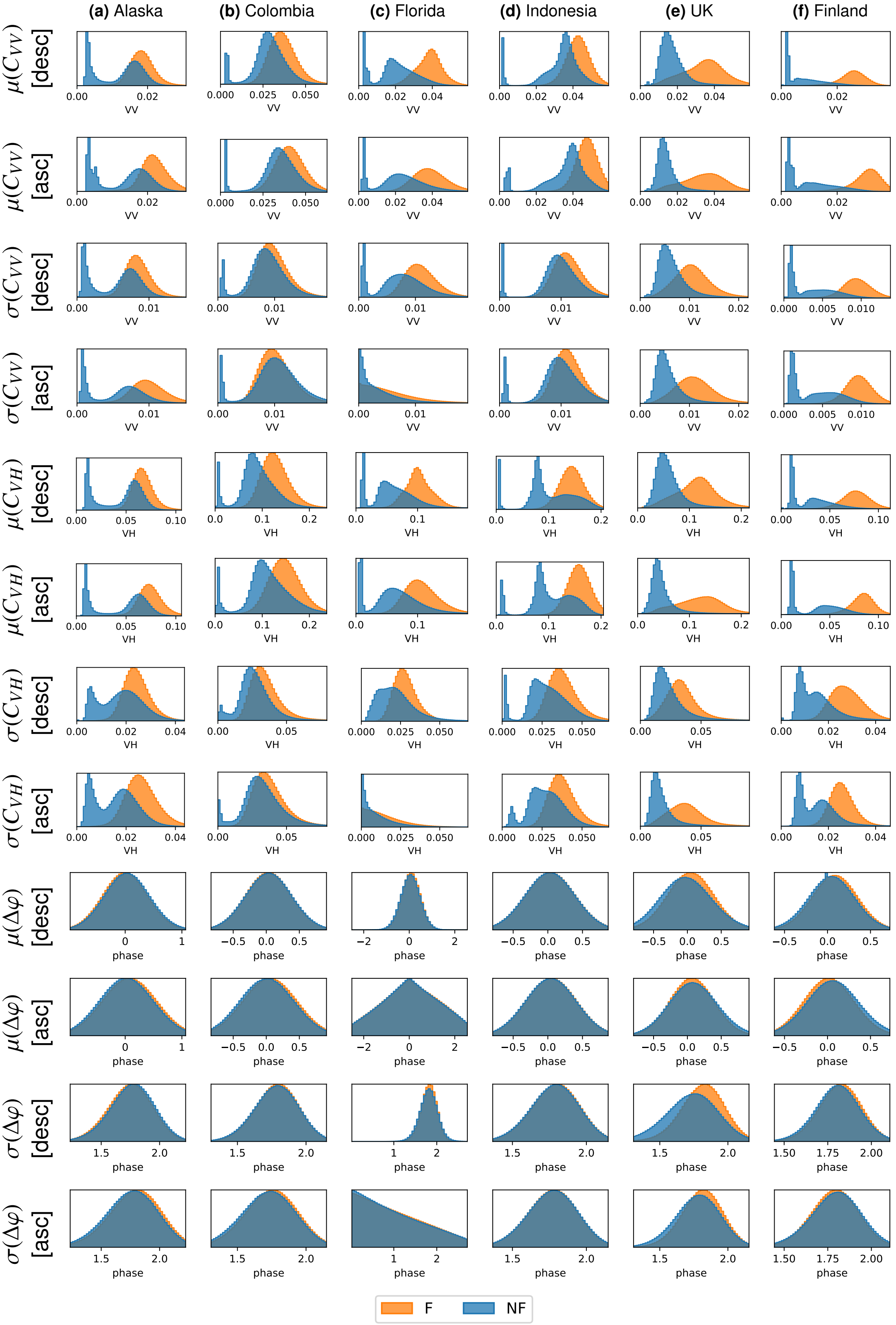


**Figure 3.5:** This figure shows the distribution of the annual mean and standard deviation of VV and VH backscatter as well as the phase difference per land cover class (forest, non-forest) according to the reference data. Forest distributions are shown in orange, non-forest in blue. Each row shows either mean or standard deviation for one of the variables, for one orbit direction. The columns represent the distributions for the different study sites.

In addition to the univariate distributions, Figure 3.6 shows two-dimensional Kernel Density Estimation (KDE) plots for pairs of variables. Each subfigure shows the distribution of $C_{VV}$ and $C_{VH}$ split between forest (in orange) and non-forest (in blue). The first two rows show the annual mean ($\mu$), whereas the last two rows show the standard deviation ($\sigma$), each for ascending and descending orbits. The shade represents the observation density, with darker shades indicating higher density.

These plots highlight the separation of forest and non-forest in a higher-dimensional feature space. Just like for the one-dimensional plots, it is clear that some sites such as Finland separate very well, whereas other sites like Colombia have more ambiguous distributions. Overall, there is a high correlation between $C_{VV}$ and $C_{VH}$ both in mean and standard deviation. This correlation is apparent from the strongly elongated distributions, especially for the mean in the UK, Finland, and Colombia. The distributions of the standard deviation show a bigger spread and separate less well, in general. In sites with very little seasonality (Colombia and Indonesia), the standard deviation appears very similar between forest and non-forest, indicating that the difference in backscatter variability is largely due to seasonal variations.

There were very few images available for the ascending orbit and the chosen year over the Florida site, so the two subfigures depicting mean and standard deviation for this orbit direction may be misleading and should be ignored.

Figure 3.7 depicts a more detailed analysis of the bivariate data distributions for the example of the Finnish site. Each subfigure marks one combination of two of the eight features (excluding the phase), where the diagonal subfigures show the corresponding one-dimensional distributions already seen in Figure 3.5. This plot unveils a strong correlation between all of the variables and as such indicates a possibility of dimensionality reduction. It is also apparent that (at least for this particular site), $C_{VV}$ tends to be a slightly better predictor of forest/non-forest than $C_{VH}$, and similarly the mean appears more informative than the standard deviation. Non-forest pixels exhibit a stronger correlation between variables, especially between the two orbit directions.

In order to quantify the separability of forest and non-forest for each site, Table 3.2 lists the overlap coefficients for each variable and site. The overlaps were derived for the temporal mean and standard deviation, as well as an individual image ($t = 0$). In addition, a multidimensional overlap coefficient was computed for each of the groups of features ($t = 0$, mean, and standard deviation) and is labeled as

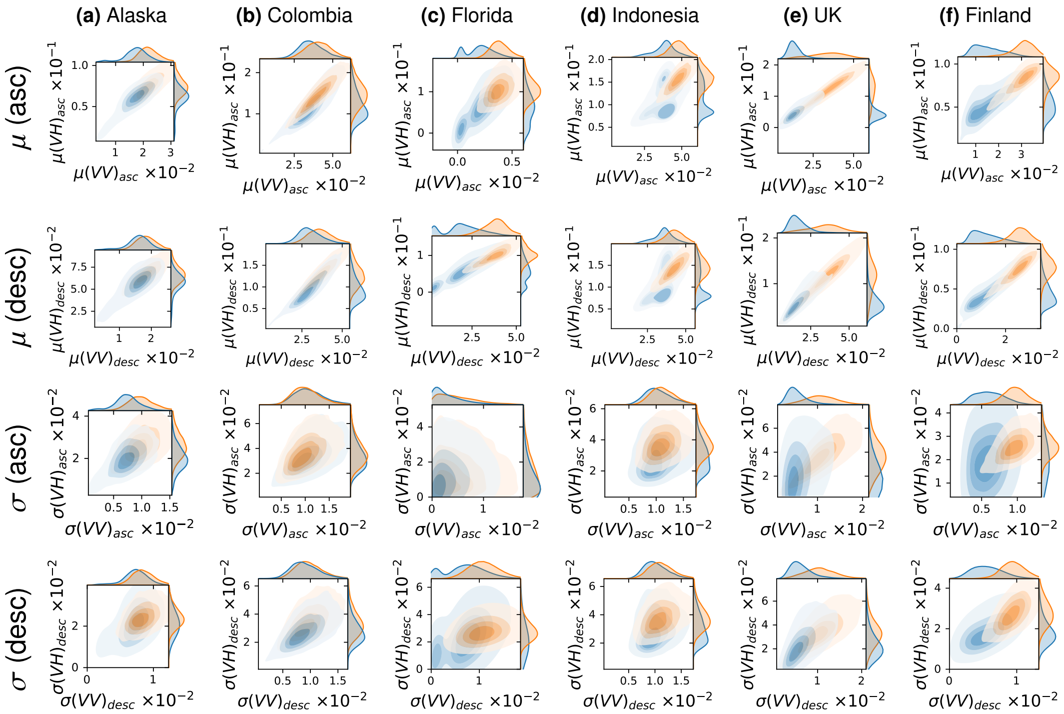


**Figure 3.6:** The subplots in this figure show the pairwise distributions of the annual mean and standard deviation of $C_{VV}$ and $C_{VH}$, for ascending and descending orbits. The distributions of forest pixels are shown in orange, non-forest in blue. Darker shades indicate a higher observation density.

"*all*" in the table. The rows labeled "*all (no* $\Delta\varphi$*)*" consider only the intensity of VV and VH backscatter, but not the phase. The overlap coefficient for combined mean and standard deviation was not computed due to computational limitations, as a 12-dimensional histogram of even as little as 20 bins per dimension would require around 100 Petabytes of memory.

The values show what was already apparent from the histograms: the separability of forest and non-forest varies to a large extent between different study sites, with the UK site appearing generally the most separable and the Alaskan site the least separable. Also, the phase difference shows upwards of 90% overlap in almost all cases except for the UK, indicating little value for telling forested areas from non-forested areas.

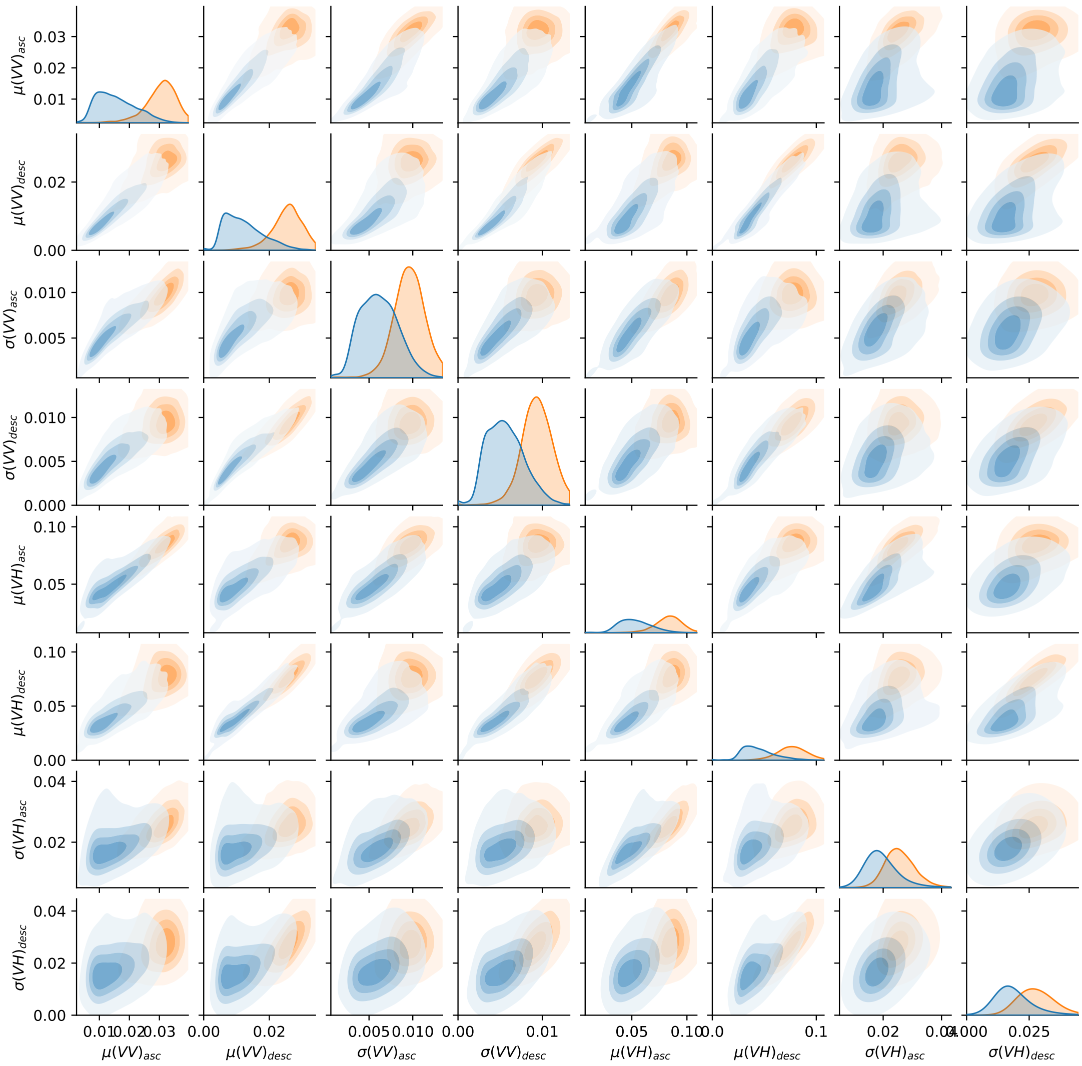


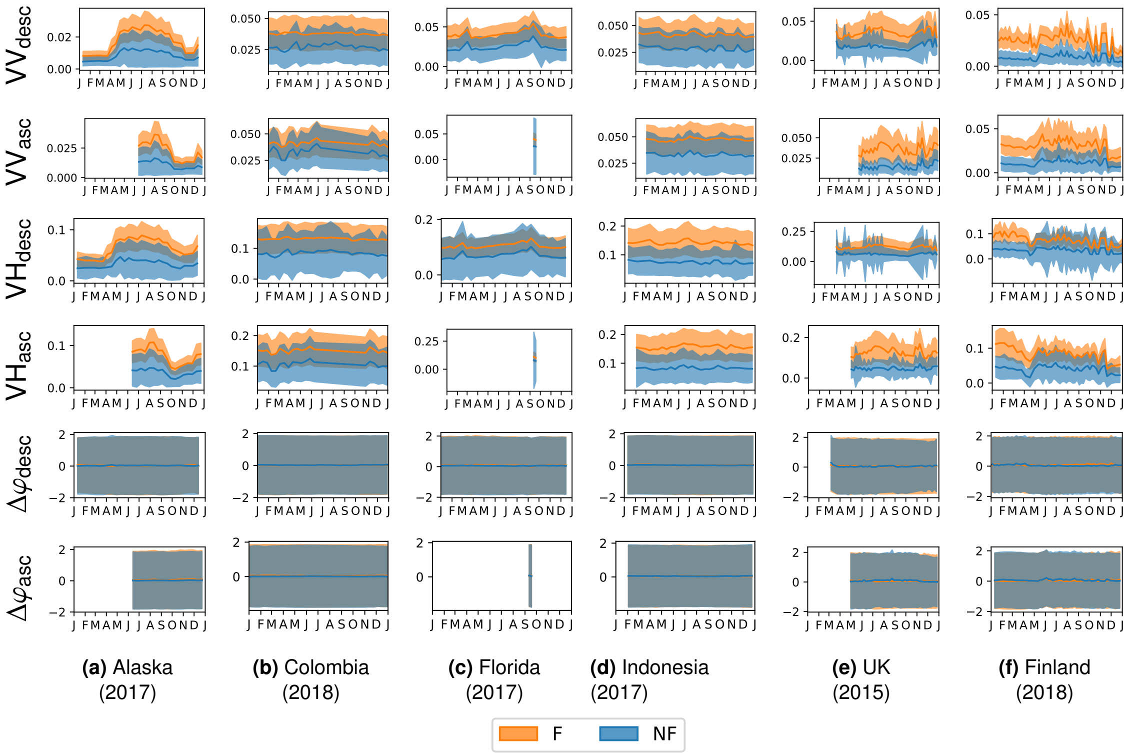
**Figure 3.7:** This grid of figures shows the bivariate distribution of the combination of several temporal statistics for the site in Finland. The subfigures on the diagonal show the univariate distributions of the respective variable. Forest distributions are shown in orange, whereas non-forest is shown in blue.

**Table 3.2:** This table shows the overlap coefficients between the forest and non-forest distributions for a variety of different variables including $C_{VV}$ (VV backscatter), $C_{VH}$ (VH backscatter), and the phase difference $\Delta\varphi$, each for ascending and descending orbit, as well as combinations thereof. The three horizontal sections show the results for a single image ($t = 0$), as well as the mean ($\mu$) and the standard deviation ($\sigma$). A value close to 0 means low overlap and high separability, whereas a value close to 1 means large overlap and low separability.

| | **OVL** | Alaska | Colombia | Florida | Indonesia | UK | Finland |
|---|---|---|---|---|---|---|---|
| $t = 0$ | $C_{VV,\text{asc}}$ | 0.457 | 0.691 | 0.325 | 0.496 | 0.224 | 0.133 |
| | $C_{VV,\text{desc}}$ | 0.546 | 0.598 | 0.239 | 0.674 | 0.385 | 0.131 |
| | $C_{VH,\text{asc}}$ | 0.468 | 0.548 | 0.353 | 0.350 | 0.244 | 0.152 |
| | $C_{VH,\text{desc}}$ | 0.518 | 0.473 | 0.254 | 0.522 | 0.502 | 0.162 |
| | $\Delta\varphi_{\text{asc}}$ | 0.974 | 0.950 | 0.988 | 0.979 | 0.960 | 0.949 |
| | $\Delta\varphi_{\text{desc}}$ | 0.920 | 0.988 | 0.974 | 0.994 | 0.780 | 0.971 |
| | all | 0.181 | 0.113 | 0.112 | 0.171 | 0.060 | 0.060 |
| | all (no $\Delta\varphi$) | 0.441 | 0.373 | 0.191 | 0.296 | 0.152 | 0.110 |
| $\mu$ | $\mu(C_{VV})_{\text{asc}}$ | 0.286 | 0.597 | 0.265 | 0.282 | 0.137 | 0.094 |
| | $\mu(C_{VV})_{\text{desc}}$ | 0.423 | 0.476 | 0.189 | 0.352 | 0.190 | 0.093 |
| | $\mu(C_{VH})_{\text{asc}}$ | 0.316 | 0.427 | 0.301 | 0.254 | 0.147 | 0.112 |
| | $\mu(C_{VH})_{\text{desc}}$ | 0.380 | 0.351 | 0.232 | 0.306 | 0.216 | 0.110 |
| | $\mu(\Delta\varphi)_{\text{asc}}$ | 0.922 | 0.924 | 0.989 | 0.991 | 0.918 | 0.883 |
| | $\mu(\Delta\varphi)_{\text{desc}}$ | 0.944 | 0.987 | 0.953 | 0.994 | 0.840 | 0.900 |
| | all | 0.204 | 0.175 | 0.123 | 0.104 | 0.073 | 0.069 |
| | all (no $\Delta\varphi$) | 0.250 | 0.236 | 0.162 | 0.118 | 0.107 | 0.086 |
| $\sigma$ | $\sigma(C_{VV})_{\text{asc}}$ | 0.305 | 0.800 | 0.621 | 0.645 | 0.221 | 0.145 |
| | $\sigma(C_{VV})_{\text{desc}}$ | 0.428 | 0.749 | 0.403 | 0.635 | 0.294 | 0.133 |
| | $\sigma(C_{VH})_{\text{asc}}$ | 0.323 | 0.727 | 0.647 | 0.400 | 0.215 | 0.216 |
| | $\sigma(C_{VH})_{\text{desc}}$ | 0.436 | 0.552 | 0.455 | 0.428 | 0.400 | 0.189 |
| | $\sigma(\Delta\varphi)_{\text{asc}}$ | 0.908 | 0.923 | 0.971 | 0.990 | 0.851 | 0.933 |
| | $\sigma(\Delta\varphi)_{\text{desc}}$ | 0.935 | 0.972 | 0.916 | 0.980 | 0.680 | 0.912 |
| | all | 0.163 | 0.203 | 0.235 | 0.227 | 0.077 | 0.073 |
| | all (no $\Delta\varphi$) | 0.264 | 0.444 | 0.363 | 0.310 | 0.143 | 0.105 |

### 3.3.2 Time series

Figure 3.8 shows the mean and standard deviation for the forest and non-forest classes over time for the duration of the year, split by orbit direction. The means are indicated as solid black lines, whereas the standard deviation is shown as an uncertainty band of $\pm\sigma$ around the mean. As is apparent from the plots, Alaska and especially Florida have little data available for the ascending orbit. For the UK site, there are no data available prior to March 2015.



**(a)** Alaska (2017)
**(b)** Colombia (2018)
**(c)** Florida (2017)
**(d)** Indonesia (2017)
**(e)** UK (2015)
**(f)** Finland (2018)

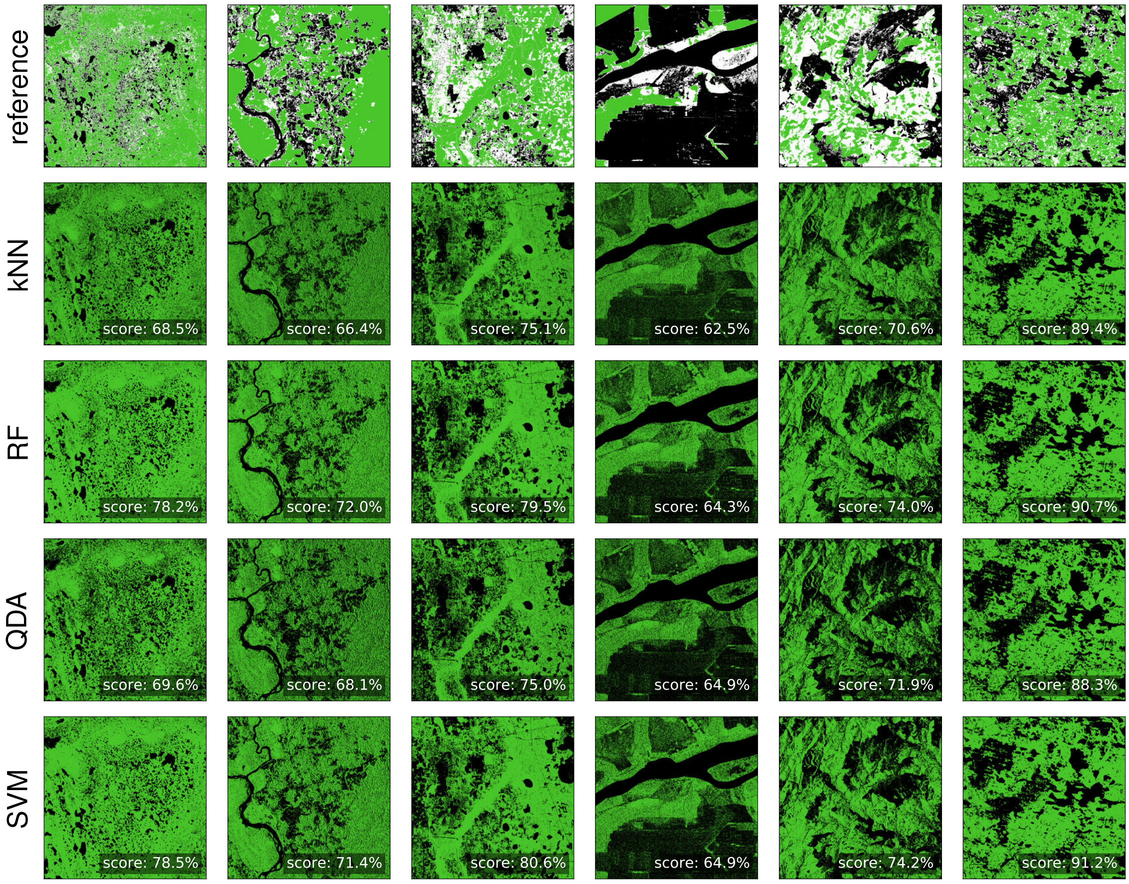
**Figure 3.8:** This figure shows the time series for forest and non-forest with standard deviation indicated as bands.

In this representation of the data it appears that the two classes do not in fact separate particularly well, as underlined by the fact that for the most part the means of one class lie within one standard deviation of the mean of the other class.

### 3.3.3 Classification results

To see how the theoretical separability holds up in practice, a number of classifiers were trained on the labeled data and their results and scores are compared in this section. The classifiers tested were kNN, Random Forest (RF), QDA, and SVM.

Figure 3.9 shows the outcome of these classifiers trained on a single time step, corresponding to case 1 in the list of feature sets in section 3.2.4. The results from using the whole time series (case 2) and the annual mean and standard deviation (case 3(a)) are shown in fig. 3.10 and fig. 3.11, respectively. The prediction images are given alongside the reference data, with forest pixels shown in green and non-forest pixels in black.

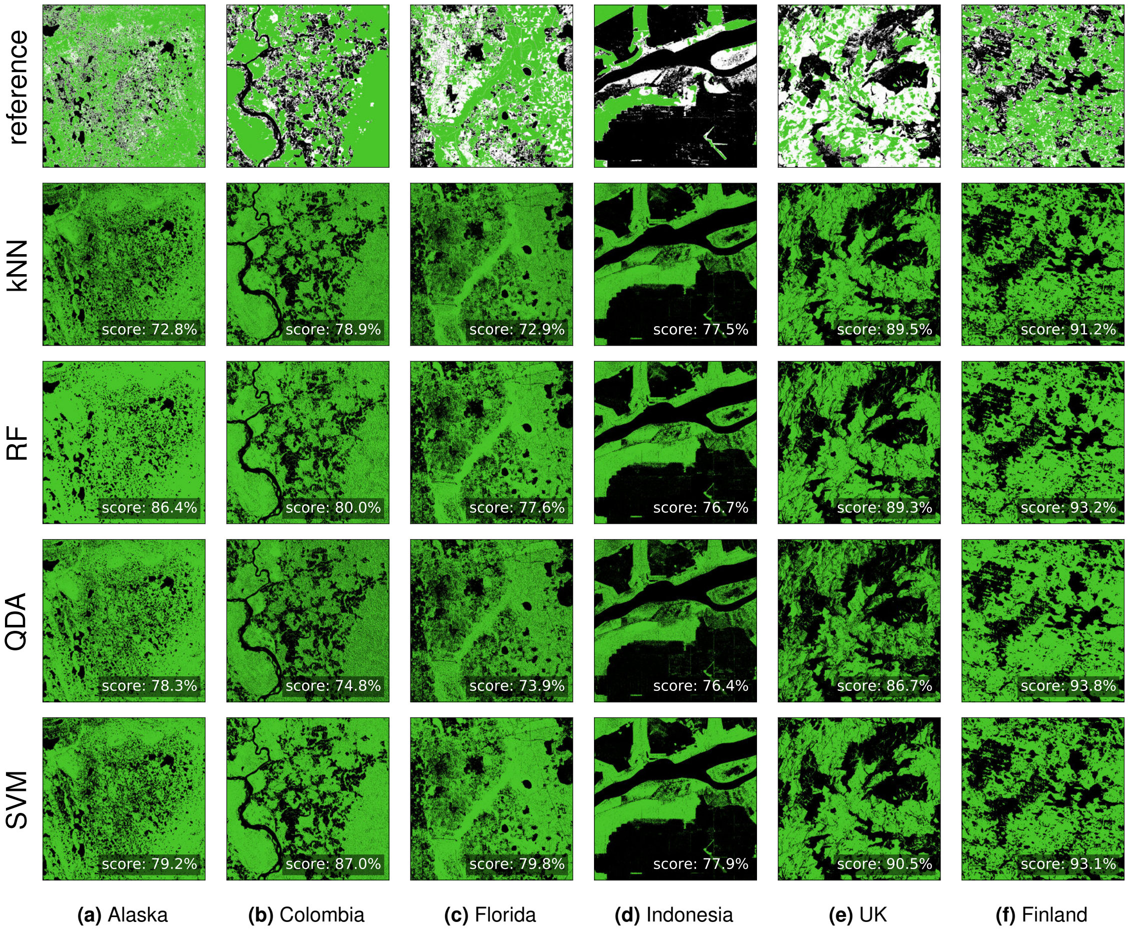


**(a)** Alaska **(b)** Colombia **(c)** Florida **(d)** Indonesia **(e)** UK **(f)** Finland

**Figure 3.9:** This figure shows classification results with different classifiers. For each study site, 1000 points were selected at random for training the classifier. Each image also shows the classification score compared to the reference data.

Random forest and SVM classifiers are favorites among the remote sensing community and indeed they appear to give the best scores overall. The kNN results in particular appear very noisy; this is because kNN is not very robust in the presence of noisy variables like the phase and in fact the results of the kNN classification improve when removing the phase based features. Table 3.3 collates the balanced accuracy scores for each of the aforementioned feature sets.

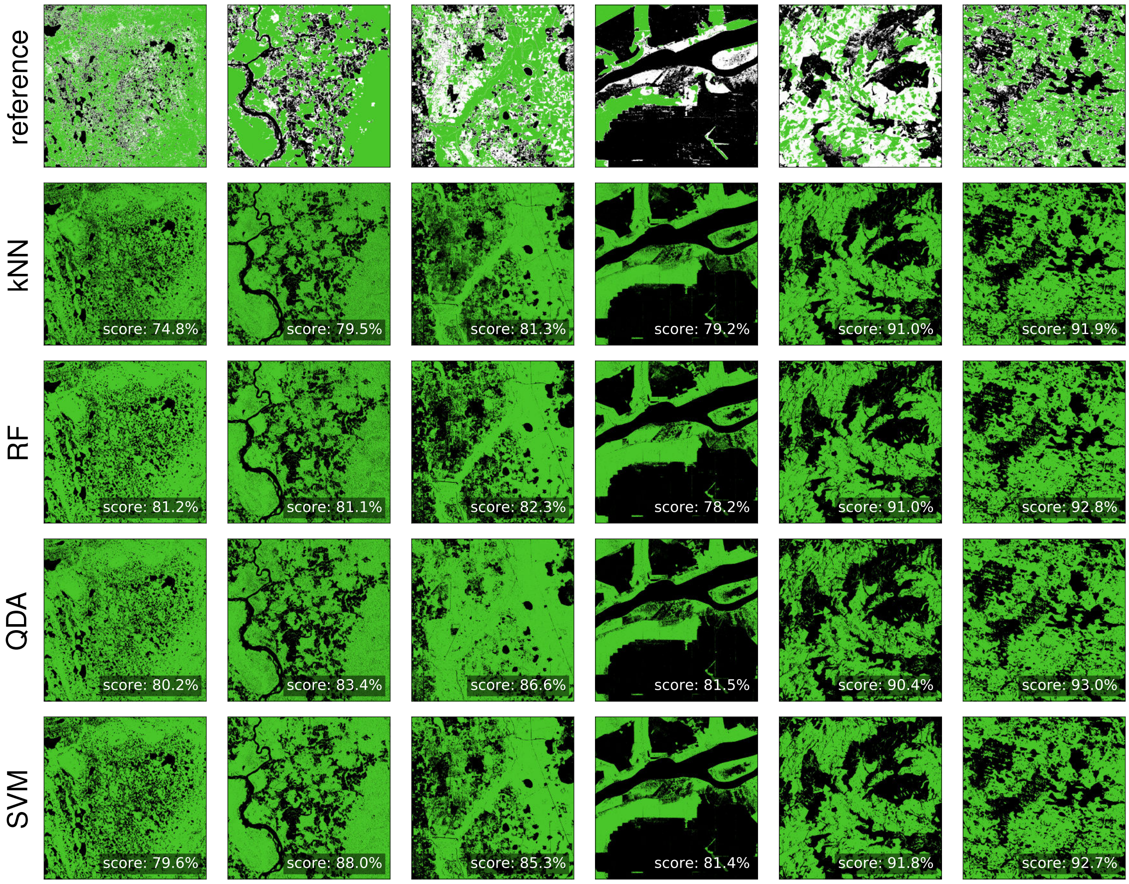
For classification using a single time step, the balanced scores are generally around 65% to 80%, varying strongly between sites. For Indonesia, the classification



**(a)** Alaska **(b)** Colombia **(c)** Florida **(d)** Indonesia **(e)** UK **(f)** Finland

**Figure 3.10:** This figure shows the classification outcome like in fig. 3.9, except this time the whole time series was used as feature vector for training and prediction.

scores are less than 65%, whereas in Finland, scores of more than 90% are observed even for a single time step. For most sites, the scores improve drastically when considering the entire time series. That is to be expected as is mitigates variations in backscatter not attributable to land cover, such as topography, seasonality, and random fluctuations. In general, the scores improve to around 75% to 90% when using the full time series as feature vector. There is a further small increase in accuracy when only the mean and standard deviation are used as features. For Florida, some of the scores decrease when the full time series data are used for classification, which could indicate that the reference land cover was collected in the beginning of the year and has subsequently changed, such that the data at $t = 0$ may be the best predictor for land cover. In the cases presented here, SVM generally performs best, whereas kNN yields the lowest scores. The exception is Finland, where there is sufficient class separation for the QDA to outperform other classifiers due to its robustness to noise.

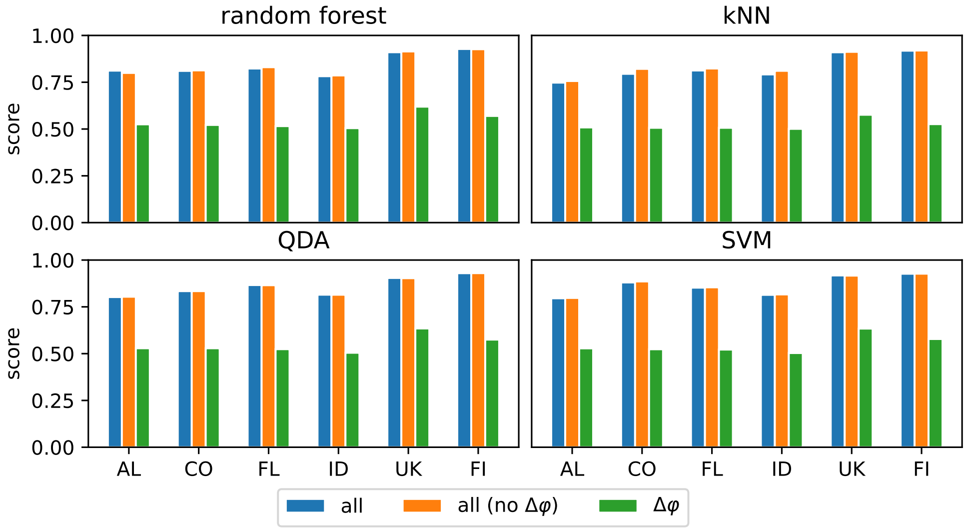


**(a)** Alaska **(b)** Colombia **(c)** Florida **(d)** Indonesia **(e)** UK **(f)** Finland

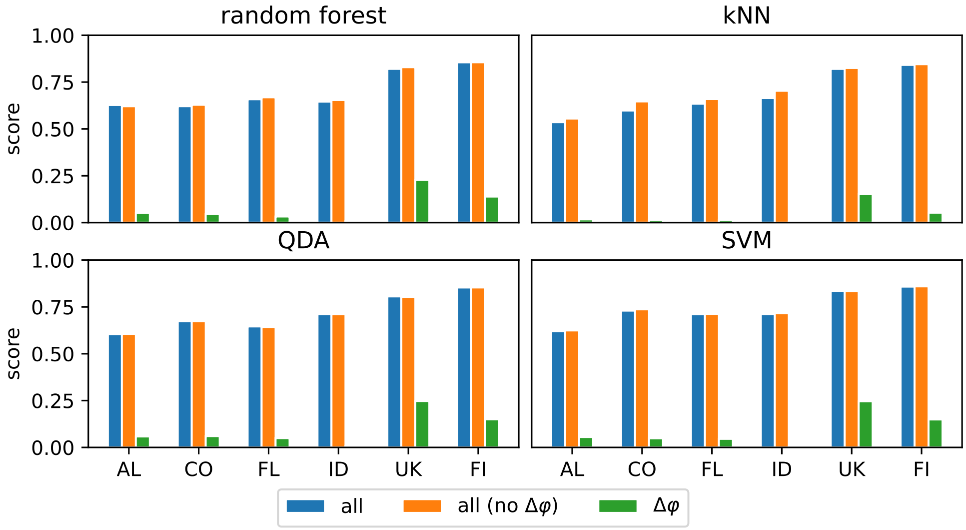
**Figure 3.11:** This figure shows the classification outcome like in fig. 3.9, except this time the mean and standard deviation (per orbit) over the whole time series were used as features for training and prediction.

Figure 3.12a shows the balanced accuracy scores for each of the classifiers and study sites. Similarly, fig. 3.12b shows Cohen's Kappa scores instead. The three different cases presented here correspond to the last three sections of table 3.3.

These plots serve to highlight the poor performance of phase-based features for classifying forest and non-forest: When training the classifiers using the phase exclusively, the balanced accuracy scores are barely above $50\%$ indicating no significant improvement over random guessing. This shows even more clearly in the Kappa scores which are all near $0\%$, with the exception of the UK and Finland where we have already seen a slight separation between forest and non-forest in fig. 3.5. Even for these two sites, however, there is no improvement when including the phase in addition to $C_{VV}$ and $C_{VH}$. We can finally conclude that the phase alone contains almost no useful information for discriminating between forest and non-forest as was already apparent from the high overlap coefficient between the phase distributions of the two classes. In most cases, the classification results are



**(a)** Balanced accuracy scores



**(b)** Kappa scores

**Figure 3.12:** This figure visualizes the performance of the four classifiers tested (random forest, kNN, QDA, and SVM) for each study site, as measured by the balanced accuracy score (**a**) and Cohen's Kappa (**b**). Three different subsets of variables are represented; VV, VH, and the phase difference $\Delta\varphi$ (blue), only VV and VH (orange), and only $\Delta\varphi$ (green).

equally good or even better when excluding the phase. This is true in particular for kNN, which does not work well in the presence of noise variables (the phase in this case). Only for the simplest model (QDA), and to a lesser extent for random forest and SVM, does the presence of the phase not affect the classifier performance. Overall, including phase information in the feature set may in fact reduce classifier performance.

**Table 3.3:** This table summarizes the balanced accuracy scores from fig. 3.11 and equivalent outcomes from using a single time step or the entire time series. The variables used for classification are VV and VH backscatter as well as the VV-VH phase difference ($\Delta\varphi$). The rows differ in whether these variables were considered for a single time step ($t = 0$), the entire time series, or the mean and standard deviation ($\mu$ and $\sigma$) of the time series, and whether or not they include the phase difference.

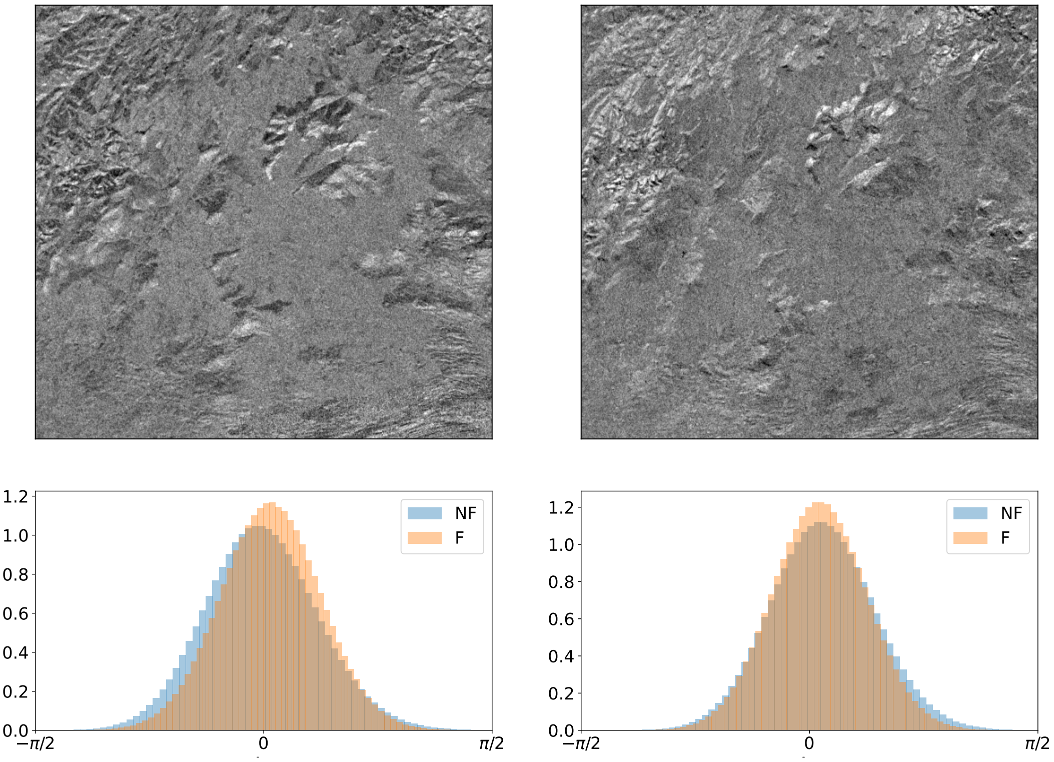
| | **Scores** | Alaska | Colombia | Florida | Indonesia | UK | Finland | *Mean* |
|---|---|---|---|---|---|---|---|---|
| $t = 0$ | kNN | 68.5% | 66.4% | 75.1% | 62.5% | 70.6% | 89.4% | *72.1%* |
| | RF | 78.2% | 72.0% | 79.5% | 64.3% | 74.0% | 90.7% | *76.5%* |
| | QDA | 69.6% | 68.1% | 75.0% | 64.9% | 71.9% | 88.3% | *73.0%* |
| | SVM | 78.5% | 71.4% | 80.6% | 64.9% | 74.2% | 91.2% | *76.8%* |
| time series | kNN | 72.8% | 78.9% | 72.9% | 77.5% | 89.5% | 91.2% | *80.5%* |
| | RF | 86.4% | 80.0% | 77.6% | 76.7% | 89.3% | 93.2% | *83.9%* |
| | QDA | 78.3% | 74.8% | 73.9% | 76.4% | 86.7% | 93.8% | *80.6%* |
| | SVM | 79.2% | 87.0% | 79.8% | 77.9% | 90.5% | 93.1% | *84.6%* |
| $\mu$ and $\sigma$ (all) | kNN | 74.8% | 79.5% | 81.3% | 79.2% | 91.0% | 91.9% | *82.9%* |
| | RF | 81.2% | 81.1% | 82.3% | 78.2% | 91.0% | 92.8% | *84.4%* |
| | QDA | 80.2% | 83.4% | 86.6% | 81.5% | 90.4% | 93.0% | *85.9%* |
| | SVM | 79.6% | 88.0% | 85.3% | 81.4% | 91.8% | 92.7% | *86.5%* |
| $\mu$ and $\sigma$ (no $\Delta\varphi$) | kNN | 75.7% | 82.1% | 82.3% | 81.1% | 91.2% | 92.0% | *84.1%* |
| | RF | 80.0% | 81.3% | 83.0% | 78.6% | 91.5% | 92.6% | *84.5%* |
| | QDA | 80.4% | 83.4% | 86.6% | 81.5% | 90.3% | 93.0% | *85.9%* |
| | SVM | 79.8% | 88.6% | 85.4% | 81.6% | 91.7% | 92.7% | *86.6%* |
| $\mu$ and $\sigma$ ($\Delta\varphi$ only) | kNN | 50.9% | 50.7% | 50.7% | 50.1% | 57.6% | 52.6% | *52.1%* |
| | RF | 52.5% | 52.2% | 51.6% | 50.5% | 62.0% | 57.0% | *54.3%* |
| | QDA | 52.9% | 52.9% | 52.4% | 50.5% | 63.5% | 57.5% | *54.9%* |
| | SVM | 52.8% | 52.3% | 52.2% | 50.3% | 63.4% | 57.8% | *54.8%* |

### 3.3.4 UK phase anomaly

We now analyze the correlation between phase difference and forest cover that was apparent in the UK study site but not—or to a much lower degree—in the other sites.

Figure 3.13 shows the spatial distribution of the VV-VH phase difference. There is a clear spatial structure both for the ascending and descending orbit. The pattern suggests a relation between the phase and the topography, specifically the slope orientation of the terrain.

To verify this suspected relation, we computed the correlation between the phase difference as well as various terrain parameters. The platform heading of the satellite is about $-16^\circ$ for the ascending orbit and $-163^\circ$ for the descending orbit, relative to north. Figure 3.14 shows hillshade images generated from a DEM of the study



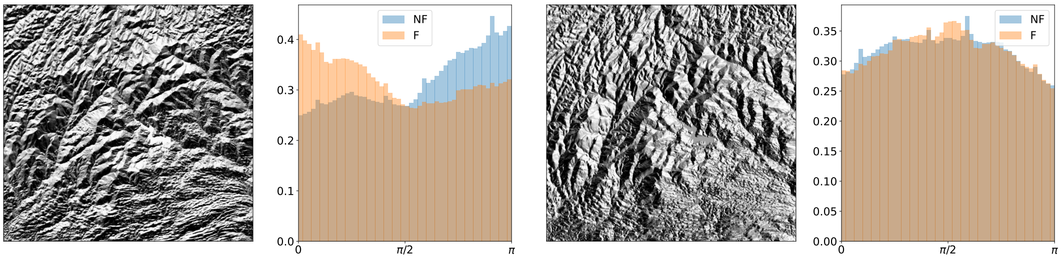
**(a)** descending orbit

**(b)** ascending orbit

**Figure 3.13:** This figure shows the distribution of the VV-VH phase difference ($\Delta\varphi$) in descending (**a**) and ascending orbit (**b**) for the UK study site. The top row shows the phase difference for each pixel, while the bottom row shows the corresponding distribution separately for forest and non-forest.

site. The hillshade illumination intensity $I$ is computed from the DEM as a function of the aspect of the terrain slope $\alpha$ and the azimuth angle $\theta$ as

$$I_\theta \propto \cos\left(\pi - \theta - \alpha\right). \tag{3.11}$$



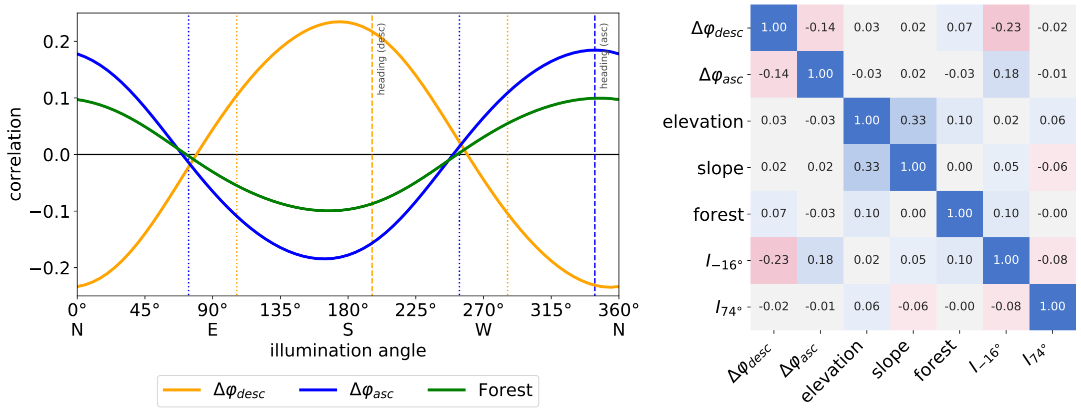
**(a)** $-16$ degrees

**(b)** $-16$ degrees

**(c)** 74 degrees

**(d)** 74 degrees

**Figure 3.14:** (**a**,**c**) depict hillshade images for illumination (azimuth) angles of $-16°$ and $74°$, respectively. This corresponds to the perceived illumination intensity for a source pointed in azimuth direction (**a**) of the ascending orbit and in perpendicular orientation (**c**). The histograms to the right of the hillshade images show the distribution of this illumination intensity separately for forest and non-forest.



| | $\Delta\varphi_{desc}$ | $\Delta\varphi_{asc}$ | elevation | slope | forest | $I_{-16°}$ | $I_{74°}$ |
|---|---|---|---|---|---|---|---|
| $\Delta\varphi_{desc}$ | 1.00 | -0.14 | 0.03 | 0.02 | 0.07 | -0.23 | -0.02 |
| $\Delta\varphi_{asc}$ | -0.14 | 1.00 | -0.03 | 0.02 | -0.03 | 0.18 | -0.01 |
| elevation | 0.03 | -0.03 | 1.00 | 0.33 | 0.10 | 0.02 | 0.06 |
| slope | 0.02 | 0.02 | 0.33 | 1.00 | 0.00 | 0.05 | -0.06 |
| forest | 0.07 | -0.03 | 0.10 | 0.00 | 1.00 | 0.10 | -0.00 |
| $I_{-16°}$ | -0.23 | 0.18 | 0.02 | 0.05 | 0.10 | 1.00 | -0.08 |
| $I_{74°}$ | -0.02 | -0.01 | 0.06 | -0.06 | -0.00 | -0.08 | 1.00 |

**(a)** **(b)**

**Figure 3.15:** The figure on the left (**a**) shows the correlation between a simulated hillshade image of the UK study site and the mean VV-VH phase difference in ascending and descending orbit as well as the forest cover for different illumination angles. Dashed and dotted lines indicate the heading angle and perpendicular directions, respectively. The figure on the right (**b**) shows the correlation between different variables in matrix form. $I_\theta$ denotes the hillshade with illumination angle $\theta$.

To illustrate the dependence of the observed correlation effect on the azimuth angle, the figure shows one hillshade image for $-16°$, i.e., the heading angle in the ascending orbit, and another for $74°$, i.e., for an illumination source displaced by $90°$. The histograms to the right of the images show the corresponding distributions of the simulated illumination intensity for forest and non-forest for both angles.

The histograms clearly show that there is a correlation effect between forest cover and terrain slope orientation that disappears when the illumination source is rotated by $90°$. Figure 3.15 quantifies this correlation and its dependence on the azimuth angle: fig. 3.15a shows the correlation of the mean VV-VH phase difference in ascending and descending orbit as well as the forest cover with the hillshade image for different azimuth angles $\theta$. The dashed and dotted lines indicate the heading angle and perpendicular directions, respectively. For the ascending orbit, these lines correspond to the special cases shown in fig. 3.14. Figure 3.15b shows the correlation with additional terrain parameters such as elevation and slope in matrix form. There is a mild correlation between the hillshade illumination at $-16°$ and the phase difference of -0.23 in the descending orbit, and 0.18 in the ascending orbit. Combined with the observed relation between forest cover and the hillshade illumination (correlation of 0.1), this is sufficient to explain the small correlation between forest cover and phase difference in descending orbit of 0.07. Albeit small, this correlation leads to the shift between forest and non-forest distributions shown in fig. 3.13a.

As the phase is a geometric phenomenon it is expected to exhibit a dependence on the ground geometry, i.e., the topography. The correlation between hillshade and forest cover is surprising, however. One possible explanation is the difference in solar exposure of north-facing vs. south-facing slopes which could drive differences in vegetation growth. However, this explanation would lead to expect maximum correlation at an azimuth angle of exactly $0°$. Instead, the observed correlation appears to correspond to the actual heading angle of the satellite. Other site-specific differences between north- and south-facing slopes are possible. Another explanation could be an intrinsic bias in the reference data. This is unlikely, though, as all reference data have been derived exclusively from optical data and are thus not affected by the same geometric considerations.

We can finally conclude that the apparent correlation between phase difference and forest cover is likely due to a mutual correlation with the terrain slope angle. After accounting for this effect the phase difference shows little prediction power for forest cover.

## 3.4 Discussion

The presented results suggest that separation between forest and non-forest is indeed possible using Sentinel-1 C-band data. In general, the separability of these two classes was greatly dependent on the study site. A larger selection of study sites may be required to determine whether this is due to the characteristics of the individual study sites or a feature of the broader climate zones, forest types, and types of forest disturbance found in each site. Overall, forest was characterized by larger backscatter values in both VV and VH, as well as larger annual variation in VV and VH compared to non-forest. This is in contrast to the results by Quegan, Toan, et al. (2000), who found that forest exhibits a *lower* temporal variability than non-forest. This may be explained by two factors. First, they mainly contrast forest with agriculture, which has a strong annual variability in backscatter due to growth and harvest cycles. Second, the low variability in backscatter was primarily found in older forests, where large biomass values are likely to be above the saturation point for biomass–backscatter relationships, and no backscatter is likely to reach the soil through dense canopies so differences in soil moisture will not change the backscatter observed.

While it was possible to achieve some separation between forest and non-forest even with a single image, using a time series or statistics from a time series was clearly helpful in all cases. Balanced accuracy scores of around 80% to 93% were possible across all study sites when data from multiple scenes were used, but the best accuracy ranged from just 65% up to 91% when a single scene was used. We found no benefit to using the full time series over the simple temporal statistics of mean and standard deviation. In fact, there was a significant increase in classification accuracy when reducing the feature set in this way, due to the notable dimensionality reduction and noise removal.

All classifier scores presented here should not be taken as the best achievable classification results. The goal of this study was not to find the best possible classifier, and as such no efforts were undertaken to optimize any hyperparameters or tune the classifiers. The emphasis was instead on identifying trends in classifier behavior when presented with different sets of features derived from Sentinel-1 SLC data. Furthermore, the scores assume that the reference labels are the “ground truth”, i.e., do not carry an error in themselves. As shown in fig. 3.3 and table 3.1, there exists a discrepancy between different references and as such these references are unlikely to be error free. The classification scores therefore do not necessarily reflect on the separability of the forest and non-forest classes alone, but could also be driven by mislabelled data.

Nevertheless, we can make some generalized observations regarding the relative performance of the four classifiers tested. When the classifiers are trained on the full time series, i.e., in the presence of high-dimensional data and high noise, an SVM or random forest classifier was found to give the best result. For a reduced feature set, SVM marginally outperforms random forest, on average.

The performance of the kNN classifier depends greatly on appropriate feature selection. While for the smallest feature set—consisting of annual mean and standard deviation of backscatter alone—this classifier showed similar performance to the other classifiers, adding noisy variables such as the cross-polarized phase difference reduced its mean accuracy from 84.1% to 82.9%. In the presence of high-dimensional data such as the full time series, its accuracy was further reduced to 80.5%. Given that kNN also does not scale well with large numbers of observations (the algorithm has to search the entire training data for each new data point) we would not recommend using it for forest classification based on these data.

Being the conceptually simplest classifier, QDA is fast and performs well in easy classification problems, i.e., scenarios where the classes show clear separation. This is true in particular of the Finnish study site. As this is not the case for other areas, however, we would not recommend using this classifier in general.

The accuracies achieved here are comparable to the results found in similar studies using diverse sensors. Dostálová, Hollaus, et al. (2016) use Sentinel-1 GRD data to map forest area over a site in eastern Austria. They achieve a balanced accuracy of 90% compared to validation data obtained from airborne full-waveform laser scanning. Our study complements this result by testing a variety of study sites, as well as quantifying the separation and thus relevance of multiple variables and derived temporal statistics, including phase information. In a global forest/non-forest map generated from L-band ALOS PALSAR data (Shimada, Itoh, et al. 2014), the achieved accuracy was 91.3% compared with Google Earth imagery, 94.8% compared with the FRA, and 84.9% compared with the DCP. Other studies have used optical data, such as Landsat images, to create similar maps. Mayaux et al. (2006) assess the accuracy of the Landsat-based Global Land Cover (GLC) 2000 map individually for each land cover class. By combining forest classes and non-forest classes, respectively, the resulting balanced accuracy score for forest/non-forest classification was 91.3%. In a European forest/non-forest map also based on Landsat data, Pekkarinen et al. (2009) achieve a balanced accuracy of 85.6%.

The UK study site exhibited an unexpected shift in phase distributions between forest and non-forest for one particular orbit direction. Our analysis of the topography showed that this shift is driven by the terrain slope orientation. Consequently, while the phase appeared to carry some information, for this site, this is an artifact of the topography and does not make a sufficient case for the inclusion of the phase in these classification efforts. The dependence of forest cover on the terrain is a site specific property that can not be generalized. As this phenomenon is independent of the radar data, the relationship between forest and terrain should be modeled separately rather than using the phase as a proxy for terrain in this way. These results suggest that the phase information of dual-polarized SAR data may be left out of forest monitoring applications unless it is used to derive interferometric quantities, in particular the temporal coherence. Coherence is a measure of the consistency of the phase over time. Because vegetation tends to result in lower temporal coherence (Askne and Smith 1997; Askne, Dammert, et al. 1997; Strozzi et al. 1998; Wegmüller et al. 1997; Wegmüller et al. 1995), this could be an additional feature for land cover classification. The predictive value of the Sentinel-

1 temporal coherence has only recently been investigated by Borlaf-Mena et al. (2021), who found a small increase in forest classification accuracy when including the temporal coherence alongside the backscatter. The use of spatial coherence for forest classification has already been demonstrated by Gaveau et al. (2003), but this is not possible with Sentinel-1 as there is no static companion satellite to allow for bistatic interferometric analysis. If no interferometric analysis is to be undertaken, we would recommend to work with GRD rather than SLC data for forest monitoring.

There are some limitations to the capacity for forest monitoring using Sentinel-1 data, as highlighted by the poor classification accuracy for single-scene data and the variability of the accuracy across study sites. The need for time series data can be a problem when temporal accuracy is required at sub-annual scale, e.g., for change detection. However, in the absence of time series data some of the added noise can be compensated for with spatial smoothing, at the cost of spatial resolution. Furthermore, the differences between study sites suggest that the accuracy of time series based classification is highest in areas of pronounced seasonality, as witnessed in, e.g., Finland or the UK. Nevertheless, even in areas with low seasonality such as the tropics, accuracies of around 80% to 85% paired with the independence of SAR data from cloud cover make a strong case for its use.

## 3.5 Conclusions

In our study of six partially forested areas in Alaska, Colombia, Finland, Florida, Indonesia, and the UK, we looked at the separation between forest and non-forest pixels for different feature sets derived from Sentinel-1 SLC data.

The following are the main conclusions.

(a) Using an SVM classifier, we were able to achieve balanced accuracy scores up to 93% in Finland, with a mean accuracy of 87% across all study sites.
(b) Given an annual time series of Sentinel-1 data, it is sufficient to extract the annual mean and standard deviation (separated by orbit) for forest mapping.
(c) There was little useful information in the cross-polarized phase difference, and no information in its variability. Most classifiers performed equally well or

better when excluding the phase difference from the feature set.

(d) For high-dimensional data, such as unaggregated time series data, we suggest using an SVM or random forest classifier as these appear the most robust to noise. For aggregated statistical data (such as annual mean and variance) our findings suggest that the best results may be achieved with an SVM classifier.

In conclusion, our goal was to assess the potential of C-band SAR data for forest detection. While longer wavelengths certainly offer better performance, we have observed sufficiently high classification accuracies globally to make a case for the use of Sentinel-1 data in forest monitoring applications. This case is helped by the free and open data policy as well as high revisit frequency, offering a potential for change detection applications.

# Chapter 4

# Change Detection

## 4.1 Deforestation and Forest Degradation Monitoring

A plethora of change detection algorithms have been proposed for use with time series of remote sensing data. All of these algorithms can roughly be categorized into two approaches. The first approach is to fit a time series model to a series of observations without changes, and then detect changes based on the discrepancy between new observations and model prediction. We may call this approach *model-based* or *parametric* change detection. Individual algorithms differ in their particular time series model and their threshold for detected change. The second approach is nonparametric. Instead of trying to fit a model, this group of algorithms typically attempts to detect persistent anomalies, or uses non-parametric statistical tests for the significance of deviations.

Mitchell, Rosenqvist, et al. (2017) present an overview of existing forest degradation monitoring techniques. The authors of the review discriminate between two main approaches: (a) assessment of degradation via change in canopy cover or proxies (e.g. roads, log decks), and (b) direct quantification of loss in AGB. The mapping of proxies, such as the progression of forest roads or forest fragmentation, could serve as a risk assessment for potential degradation when combined with a proximity metric. Forest road mapping was e.g. done by Laporte et al. (2007), Rauste et al. (2013), and Solberg, Astrup, et al. (2013).

In order to monitor changes in any of these metrics, a number of change detection algorithms have been developed specifically for application in remote sensing data, including BFAST, LandTrendr, and CCDC.

Vegetation indices such as NDVI and Enhanced Vegetation Index (EVI), as well as SAR backscatter are commonly used to monitor forest canopy change. Novel approaches use InSAR coupled with a DEM to precisely monitor canopy structure changes (Deutscher et al. 2013; Solberg, May, et al. 2015).

Measuring or estimating forest biomass is a challenging task, but is ultimately the goal as it enables direct quantification forest carbon stocks and greenhouse gas (GHG) emissions. Because the amount of biomass relates to the three-dimensional structure of the forest, SAR and lidar are the most promising candidates for measuring biomass since they can penetrate the canopy— assuming a sufficiently long wavelength for SAR and a high enough return energy for lidar. Large-scale lidar surveys, however, will only be feasible with the availability of space-borne lidar sensors such as NASA's GEDI. A recent feasibility study has demonstrated the potential for global wall-to-wall coverage with a spaceborne lidar system (Hancock et al. 2021). SAR has been successfully applied for biomass estimation, but it has been shown that the backscatter signal saturates quickly and hence may be more useful for use in degraded or young forests rather than dense and mature forests (Böttcher et al. 2009; Mermoz, Le Toan, et al. 2014). This problem does not exist with lidar. Fusion of SAR, optical, and lidar data may provide the key to highly accurate and precise canopy cover and sub-canopy structure measurements.

Any method required to reach operational readiness, to be applied on national level for REDD+, will have to satisfy a range of criteria including scalability, automation, and the possibility of tuning model parameters to reflect regional differences. Further research is needed in the areas of SAR-optical data fusion, generating carbon stock baselines, InSAR DEM differencing, and the integration of Earth Observation (EO) techniques with in-situ measurements to achieve these requirements.

## 4.2 Map Accuracy

Any time we produce a map (e.g., a map of detected changes or classified land cover) we are interested in a measure of the map accuracy to assess the performance of the map generating algorithm compared with other methods.

The following definitions apply to binary maps (e.g., change vs. no change, or forest vs. non-forest). Several accuracy measures are commonly used for assessing a map. All of these can be computed from the entries of a confusion matrix, i.e., the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). The Producer's Accuracy (PA) (also known as sensitivity, recall, or true positive rate) in this context is the probability that an actual change is accurately detected. It is defined as

$$\mathrm{PA} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \tag{4.1}$$

The User's Accuracy (UA) (also known as precision or positive predictive value) is defined as the probability that a change shown on the map corresponds to a true change on the ground. It is given by

$$\mathrm{UA} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} \tag{4.2}$$

Finally, the Overall Accuracy (OA) (or simply accuracy) is defined as

$$\mathrm{OA} = \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{FP}+\mathrm{TN}+\mathrm{FN}} \tag{4.3}$$

and is the overall fraction of correctly classified pixels as change or no change.

## 4.3 Literature Review

As discussed previously, change detection algorithms can be roughly categorized as model-based or non-parametric. The following two subsections give an overview of previously proposed change detection algorithms in either of these categories.

### 4.3.1 Model-based change detection algorithms

Verbesselt, Hyndman, Newnham, et al. (2010) propose the method BFAST (*Breaks for Additive Seasonal and Trend*) for structural change detection in time series. It was noted that change in a plant phenological time series may be due to (a) seasonal change, (b) gradual change, and (c) abrupt change, whereby the latter is the phenomenon to be detected. To this end, the authors assumed a parametric time series model of the form $Y_t = T_t + S_t + e_t$, where $T_t = \alpha + \beta t$ is the trend, $S_t$ is the seasonal component, and $e_t$ is noise. A change, or break point, is then defined as a time where the parameters of the fitted model change. The seasonal component was extracted using STL (Cleveland et al. 1990) and the trend was fitted to the deseasonalized data $Y_t - \hat{S}_t$ as a piecewise linear model. Changes were recorded at the corresponding boundaries. The authors adopted an iterative approach to ensure convergence of the break point locations.

In a later paper, Verbesselt, Zeileis, et al. (2012) introduce a different change detection method for time series of environmental data with trend and seasonal effects present. *Disturbance* was defined to mean time series changes that could not be explained by either trend, seasonality, noise, or artifacts such as atmospheric scatter, cloud cover, or geometric errors. The authors adopted an online monitoring approach by fitting a time series model to an undisturbed history period and subsequently checking new data values for consistency with the model. A time series model with additive linear trend and harmonic seasonality was chosen:

$$y_t = \alpha_1 + \alpha_2 t + \sum_{j=1}^{k} \gamma_j \sin\left(\frac{2\pi jt}{f} + \delta_j\right) + \varepsilon_t \tag{4.4}$$

where $\alpha_1$ is the intercept, $\alpha_2$ is the slope of a linear trend, $\gamma_1 \ldots \gamma_k$ and $\delta_1 \ldots \delta_k$ are the amplitude and phase parameters of the harmonic components, $f$ is the known frequency depening on the number of observations per year, and $\varepsilon_t$ is a Gaussian error term. The method was applied to drought detection in Somalia using MODIS NDVI data from February 2000 to July 2011. It appeared that strong seasonal variability in the data prevented the detection of less-than-severe disturbances. The authors found that their method was computationally fast, didn't require the introduction of user-defined thresholds, and was able to handle missing data. The application to deforestation detection was explicitly left outstanding.

DeVries et al. (2015) build on the research done by Verbesselt, Hyndman, Newnham, et al. (2010) and Verbesselt, Zeileis, et al. (2012), focusing on detecting changes of smaller scale and smaller magnitude, i.e. forest degradation rather than clear-felling. The study site was the Kafa Biosphere Reserve in southern Ethiopia, where the vegetation was characterized by a mixture of fragmented humid broadleaf evergreen forest and patterns of integrated cropland-forest. Data were analyzed for the period from 1999 to 2011 with an average of more than 9 images per year. Pre-processing was done using LEDAPS (Schmidt et al. 2013) and Fmask. A forest benchmark map was generated for the beginning of the monitoring period using a supervised maximum likelihood classifier in ArcGIS as well as a mask of less than 30% forest cover obtained from the MODIS VCF product (Hansen, DeFries, et al. 2003).

The authors fitted a time series model of the form

$$y_t = \alpha + \gamma \sin\left(\frac{2\pi t}{f} + \delta\right) + \varepsilon_t \tag{4.5}$$

where $\alpha$ is the intercept, $\gamma$ and $\delta$ are the amplitude and phase of the harmonic term, $f$ is the known frequency depending on the number of observations per year, and $\varepsilon_t$ is a Gaussian error term. This model notably lacks a trend and higher order harmonic terms, in order to (a) avoid over-fitting given the limited number of observations per year and (b) avoid unrealistic projections of short-term trends into the future. The deviation of the time series from the model (*change magnitude*) was then measured as $M = \mathrm{median}(y_t - \hat{y}_t)$ where $y_t$ and $\hat{y}_t$ were the actual and expected observations. That magnitude was measured as a continuous property of the time series, rather than for discrete breakpoints. Having detected possible breakpoints using BFAST, the authors then proceeded to classify the breakpoints as either (i) no change, (ii) degradation, or (iii) deforestation, using Ordinal Logistic Regression (OLR) and the magnitude as a predictor. In a combined change map the mean size of a disturbance pixel cluster was $0.6$ ha, which well represented the fact that forest change was mostly driven by small-scale agriculture in remote locations, often in incremental steps. The authors state as main sources of error (a) remnant cloud effects and (b) inaccuracies in the forest benchmark mask, e.g. due to confusion of deforestation with crop cycle changes. The overall accuracy achieved was 78%.

Hamunyela, Verbesselt, et al. (2016) propose an extension of the BFAST algorithm by using spatial context to overcome the challenge posed by seasonal changes in deforestation detection. The problem arises when attempting to detect deforestation at sub-annual scales because depending on the type of forest, strong seasonal dependence of the phenology makes it difficult to isolate deforestation effects. Typically, the seasonality is modeled explicitly and can then be subtracted from the measured variable (deseasonalization). However, this requires at the very least one full year (ideally multiple years) of training data. The authors propose a new approach based on the central assumption that pixels in spatial proximity of each other exhibit similar seasonal dynamics. Therefore, if a subset of pixels in a spatial neighborhood undergo deforestation, that change should be visible as an increased contrast between the affected pixels and their unaffected neighbors, even if otherwise masked by seasonal changes. This contrast is captured by spatially normalizing the variable of interest. Two study areas were selected to test the method: (i) a site of humid, evergreen tropical forest in Brazil that exhibits weak seasonality and where deforestation is heterogenous and largely due to progressing colonization, and (ii) a site of dry tropical forest in Bolivia with strong seasonality, where deforestation occurs in large blocks due to industrial agricultural expansion.

In a first step, the spatial normalization was done for each image individually as

$$\mathrm{sNDVI} = \frac{\mathrm{NDVI}_{\text{pixel}}}{\mathrm{NDVI}_{\text{median}}} \tag{4.6}$$

where $\mathrm{NDVI}_{\text{median}}$ is the median of all pixels within the spatial neighborhood above or equal to the 90$^{\text{th}}$ percentile of those pixels. The spatial neighborhood was set to be a rectangular window. The authors tested eight window sizes from $9 \times 9$ to $121 \times 121$ pixels, corresponding to an area of $7.29$ to $1317.69$ ha. It turned out that the optimal window size depends strongly on the deforestation pattern, as illustrated by the fact that the best result for the Bolivian site was for $334.89$ ha, while the same window size for the Brazilian site yielded the worst result.

Following the spatial normalization, changes were detected for individual pixels using BFAST. As BFAST was not developed specifically to detect deforestation but any disturbance, the authors introduced a magnitude threshold for detected changes, above which a change would be classified as disturbance. With an optimized window size and change magnitude threshold, the overall accuracy was slightly larger than the seasonal model approach for both test sites. The largest improvement was made in the detection time lag for the Bolivian forest, where deforestation was detected after two observations, compared to 15 observations for the seasonal model.

Zhu, Woodcock, and Olofsson (2012) propose a different model-based change detection algorithm, the CMFDA (*Continuous Monitoring of Forest Disturbance Algorithm*). A study area in the Savannah River Basin between Georgia and South Carolina was selected, covered mostly by evergreen forest and agriculture. All available Landsat 5 and 7 images for the study area were used. Fmask was used to remove clouds and cloud shadows, but was complemented using a further cloud masking approach based on Top Of Atmosphere (TOA) reflectance called *Tmask* (Zhu and Woodcock 2014a). The authors look at several potential forest disturbance measures, including NDVI, red and Shortwave Infrared (SWIR) bands, Normalized Burn Ratio (NBR), Wetness (from Tasseled Cap Transformation (Crist 1985; Crist and Cicone 1984)), Integrated Forest Z-score (Huang et al. 2010) (IFZ) and Disturbance Index (Healey et al. 2005; Masek et al. 2008) (DI). The following measures were tested in this study: Band 3 reflectance, Band 7 reflectance, NDVI, NBR, wetness, and a simplified DI ($B - (G + W)$, where $B$ is brightness, $G$ is greenness and $W$ is wetness).

The following surface reflectance model was fitted to each band during a history period/training data:

$$\begin{aligned} f(x) = &a_0 + \sum_{i=1}^{N} \left( a_i \cos\left(\frac{2\pi}{iT}x\right) + b_i \sin\left(\frac{2\pi}{iT}x\right)\right) \\ &+ a_{N+1} \cos\left(\frac{2\pi}{0.5T}x\right) + b_{N+1} \sin\left(\frac{2\pi}{0.5T}x\right) \end{aligned} \tag{4.7}$$

where $x$ is the day of year, $N$ is the number of years, $T$ is the number of days per year, and $a_i, b_i$ are the $2N + 3$ model parameters. The authors used the finding that forests have high NDVI values (Masek et al. 2008) and low SWIR reflectance (Huang et al. 2010; Kennedy, Cohen, et al. 2007) to derive a stable forest mask; possible forest is marked where the mean NDVI is larger than 0.6 and mean Band 7 reflectance less than 0.1. Noting that $a_2$ and $b_2$ capture inter-annual change, the stable forest mask was finally derived as all pixels that fulfill the following conditions:

$$\frac{a_{0,4} - a_{0,3}}{a_{0,4} + a_{0,3}} > 0.6 \quad \text{and} \quad a_{0,7} < 0.1 \quad \text{and} \quad \sqrt{a_{2,7}^2 + b_{2,7}^2} < 0.02 \tag{4.8}$$

where $a_{i,j}$ denotes the parameter $a_i$ for Landsat band $j$. Starting with the reflectance model and the stable forest mask, two change detection algorithm variants were investigated: single-date and multi-date. The most successful approach was the multi-date variant, where disturbance was declared for a pixel in the stable forest mask when there were three consecutive observations for which $(B - (G + W))_{\text{obs}} - (B - (G + W))_{\text{pred}} > 0.12$. The accuracy was assessed using 21 visually interpreted rectangular regions in the study area, 11 of which were used for model fitting and 10 for model validation. For the multi-date algorithm, the authors found a user's accuracy of 95.8%, a producer's accuracy of 96.7%, and a temporal accuracy of almost 94%. The authors excluded edge pixels from their accuracy calculations, such that the actual accuracy achieved is likely much lower (given that most misclassification happens at the forest edges).

Yet another model-based change detection algorithm was introduced by Zhu and Woodcock (2014b): the CCDC (*Continuous Change Detection and Classification*). All Landsat images with less than 80% cloud cover were selected (according to Fmask). For each band $i$ the following model was fitted:

$$\hat{\rho}_i(x) = a_{i,0} + a_{i,1} \cos\left(\frac{2\pi}{T} x\right) + b_{i,1} \sin\left(\frac{2\pi}{T} x\right) + c_{i,1} x \tag{4.9}$$

where $x$ is the Julian date, $T = 365$ is the number of days per year, and $a, b, c$ are the model coefficients. Change is detected if the following condition is satisfied for three time steps in a row:

$$\frac{1}{k} \sum_{i=1}^{k} \frac{\rho_i(x) - \hat{\rho}_i(x)}{3 \cdot \mathrm{RMSE}_i} > 1 \tag{4.10}$$

where $\rho$ is the observed value, $\hat{\rho}$ is the predicted value, and $\mathrm{RMSE}$ is the root mean square error of the model fit. The authors decided to normalize by three times the RMSE to filter out noise. Some additional tweaks are included to account for possible change points at the very beginning of the time series. Furthermore, a random forest classifier provides a land cover classification at each time step. The accuracy of CCDC is cited as 98% producer's accuracy and 86% user's accuracy in the study area on the US East Coast.

### 4.3.2 Nonparametric change detection

Kennedy, Yang, et al. (2010) developed an approach called LandTrendr, which segments a yearly time series of spectral indices derived from Landsat data. The algorithm iteratively models the time series by a piecewise linear fit with an increasing number of vertices (nodes). At each iteration a new vertex is added at the point with the largest deviation from the linear fit, until a preset maximum number of segments is reached. Starting from the maximally complex model, vertices are removed based on several criteria and a goodness-of-fit criterion is calculated at each step, in this case the $p$ value of the $F$ statistic. The model with the highest score is accepted. LandTrendr relies on a single spectral index and does not take into account seasonal effects. It is therefore not capable of detecting changes at the subannual scale.

Hamunyela, Reiche, et al. (2017) note that the trade-off between spatial and temporal accuracy in deforestation detection is limiting the overall accuracy of disturbance maps. If short temporal detection delays are to be achieved, this typically results in low omission error but very high commission error. In particular, the use of change magnitude thresholds is seen as problematic as it relies on the assumption that true and false detections do indeed exhibit different change magnitudes. While this assumption may hold for large scale clear-felling, small scale or scattered forest disturbances may go undetected.

The authors investigated whether including information about the spatial neighborhood of a pixel may improve the accuracy of deforestation detection while keeping the temporal delay short. The study area was the Kafa Biosphere Reserve in Ethiopia that had already been analyzed by DeVries et al. (2015); a humid Afromontane broadleaf evergreen forest with small-scale and fragmented deforestation drivers. A total of 172 Landsat images between January 2010 and June 2016 were used to derive the NDVI time series. A benchmark forest mask was generated from multispectral Landsat data for 2013 and the pixelwise NDVI time series were spatially normalized as proposed by Hamunyela, Verbesselt, et al. (2016) to reduce seasonality effects (using a $15 \times 15$ pixel window). Notably, small forest patches of less than $0.54$ ha (6 pixels) were removed from the benchmark mask.

In a first step, they flag as possible deforestation events all occurrences of at least two consecutive negative anomalies in the Landsat NDVI time series; a common approach. Secondly, a number of spatio-temporal features of the pixel neighborhood were computed and used to decide whether or not this was a valid detection. A random forest classifier was used that had been trained on a data set derived from visual interpretation of multi-spectral Landsat data. Different thresholds were used for detecting anomalies (percentiles 1 to 5), but the authors found little difference in the final result since the spatio-temporal classifier was able to undo the increased commission error of a less strict threshold. The 17 spatio-temporal features were designed to measure (i) change in spatial variability over time, (ii) spatio-temporal proximity to other anomalous pixels, (iii) spatial proximity to non-forest areas, and others such as the anomaly threshold. The most predictive features were spatial-temporal variability change, the proximity to non-forest pixels, and overall variability. Spatial and temporal accuracy were estimated using stratified sampling, with split strata for stable forest at the edge (where misclassifications are more likely) and away from the edge. The authors postulate that their space-time features are sensor-blind and in particular also work for SAR data.

Joshi, Mitchard, et al. (2015) attempt to detect forest degradation events, i.e. the reduction of forest cover and/or biomass to a lesser degree than clear felling. The data set consists of two dry season images per year from 2007 to 2010 for a region in Madre de Dios, Peru. The authors emphasize the importance of radar imagery for this particular region due to an average cloud cover of over 70%. According to the reference data set (Hansen, Potapov, et al. 2013), deforestation was happening in the region at $0.35\,\%\,\mathrm{yr}^{-1}$. Initially, all areas with $< 30\%$ forest cover were masked (Hansen, Potapov, et al. 2013). Bitemporal change maps for any pair of years in the study period were generated (using only the HV backscatter) and disturbance was detected based on a simple threshold and the consistency of change. The authors distinguish between slow-recovering and fast-recovering changes by considering the stability of a change once detected. The forest disturbance (deforestation and degradation) rate was found to be $0.78\,\%\,\mathrm{yr}^{-1}$. The algorithm accuracy was assessed using a dataset of newly built farms in the region as well as requested harvest areas (false negatives), and a manually selected area without records of logging or forest loss (false positives). 63% of the farms and 50% of the requested harvest sites were detected as forest disturbance. The false positive rate was $0.1\,\%\,\mathrm{yr}^{-1}$.

Conradsen, Nielsen, and Skriver (2016) present a change detection algorithm for time series of complex valued SAR data based on the complex Wishart distribution for the covariance matrices. $S_{rt}$ denotes the complex scattering amplitude where $r, t \in \{h, v\}$ are the receive and transmit polarization, respectively (horizontal or vertical). Reciprocity is assumed, i.e. $S_{hv} = S_{vh}$. Then the backscatter at a single pixel is fully represented by the complex target vector $\boldsymbol{s} = \begin{bmatrix} S_{hh} & S_{hv} & S_{vv} \end{bmatrix}^T$.

For multi-looked SAR data, backscatter values are averaged over $n$ pixels (to reduce speckle) and the backscatter may be represented appropriately by the (variance-)covariance matrix, which for fully polarimetric SAR data is given by

$$\langle C \rangle_{\text{full}} = \langle \boldsymbol{s}(i)\boldsymbol{s}(i)^H \rangle = \begin{bmatrix} \langle S_{hh}S_{hh}^* \rangle & \langle S_{hh}S_{hv}^* \rangle & \langle S_{hh}S_{vv}^* \rangle \\ \langle S_{hv}S_{hh}^* \rangle & \langle S_{hv}S_{hv}^* \rangle & \langle S_{hv}S_{vv}^* \rangle \\ \langle S_{vv}S_{hh}^* \rangle & \langle S_{vv}S_{hv}^* \rangle & \langle S_{vv}S_{vv}^* \rangle \end{bmatrix} \tag{4.11}$$

where $\langle \cdot \rangle$ is the ensemble average, $^*$ denotes complex conjugation, and $^H$ is Hermitian conjugation. Often, only one polarization is transmitted (e.g. horizontal), giving rise to dual polarimetric SAR data. In this case the covariance matrix is

$$\langle C \rangle_{\text{dual}} = \begin{bmatrix} \langle S_{hh} S_{hh}^* \rangle & \langle S_{hh} S_{hv}^* \rangle \\ \langle S_{hv} S_{hh}^* \rangle & \langle S_{hv} S_{hv}^* \rangle \end{bmatrix} \tag{4.12}$$

These covariance matrices follow a complex Wishart distribution as follows:

$$\boldsymbol{X}_i \sim W_C(p, n, \boldsymbol{\Sigma}_i), \quad i = 1, ..., k \tag{4.13}$$

where $p$ is the rank of $\boldsymbol{X}_i = n\langle \boldsymbol{C}_i \rangle$, $E[\boldsymbol{X}_i] = n\boldsymbol{\Sigma}_i$, and $\boldsymbol{\Sigma}_i$ is the expected value of the covariance matrix.

Based on this distribution, a test statistic was first developed for a pair of such matrices (Nielsen et al. 2007), allowing for bitemporal change detection.

For a multitemporal sequence of images, the change detection problem then becomes a test of the null hypothesis $H_0 : \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \ldots = \boldsymbol{\Sigma}_k$, i.e. whether the expected value of the backscatter remains constant. This test is a so-called omnibus test.

A test statistic for the omnibus test can be derived as

$$Q = k^{pnk} \frac{\prod_{i=1}^{k} |\boldsymbol{X}_i|^n}{|\boldsymbol{X}|^{nk}} = \left\{ k^{pk} \frac{\prod_{i=1}^{k} |\boldsymbol{X}_i|}{|\boldsymbol{X}|^k} \right\}^n \tag{4.14}$$

where $\boldsymbol{X} = \sum_{i=1}^{k} \boldsymbol{X}_i \sim W_C(p, nk, \boldsymbol{\Sigma})$ (Conradsen, Nielsen, Schou, et al. 2003). The test statistic can be translated into a probability $p(H_0)$. The hypothesis test is repeated iteratively over subsets of the time series in order to determine the actual time of change.

Because the omnibus test makes use of the full complex-valued data of the Sentinel-1 SLC product, I chose to investigate the performance of this algorithm for deforestation detection with dual-polarimetric C-band SAR data.

## 4.4 Omnibus Test Based Change Detection

Figure 4.1 shows a time series of SAR data over the study area in Rondônia, Brazil ($9.3°$S, $63.8°$W to $9.5°$S, $63.55°$W) for the year 2017. The images are based on the dual polarization complex covariance matrix (eq. (4.12)) with the following channel assignments: $C_{VV}$ in red, $C_{VH}$ in green, $C_{VV}/C_{VH}$ in blue. This corresponds to the intensities in VV and VH as well as their ratio, because the intensity is the square of the backscatter amplitude. This RGB representation is commonly used for dual polarization data, as vegetation tends to appear in green, water in blue, and built up areas in red which corresponds to our intuitive interpretation of the image.

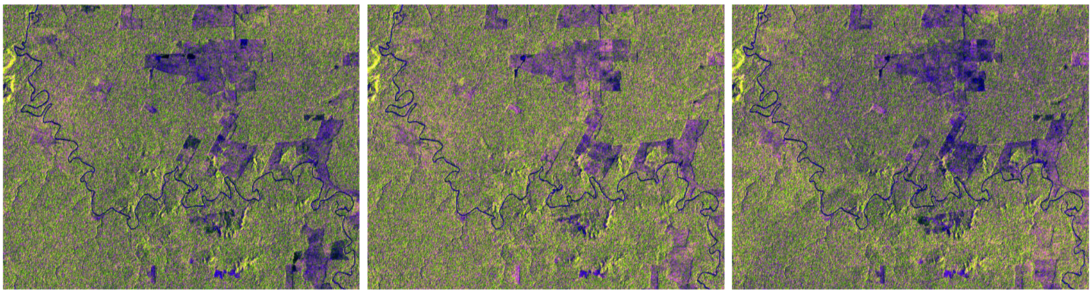
A total of 30 images (in descending orbit) are available for this area and time period. The scale is kept consistent throughout the image sequence for comparability.

Forest areas appear mostly green, at least in the images from January to June. It is also apparent how much noise (or speckle) there is in the forested area. A river is clearly visible as a sharply delineated dark blue line across the image. The darker areas with polygonal shapes are clearfellings that are being used for agriculture.

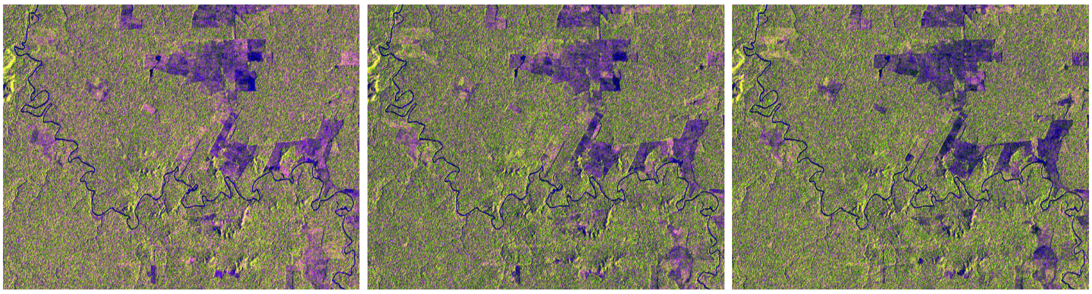
Already from a visual inspection of this time series it is clear that there are strong seasonal effects present which equally affect large portions of the image. For example, if we decide to compare only the images from January and August we may reasonably conclude that changes have occurred throughout the whole image (perhaps with the exception of water bodies). This is not inherently wrong, however, these are not the kind of changes that we are typically interested in detecting.

We can also observe radar shadow and layover effects as demonstrated in the following close up (fig. 4.2).

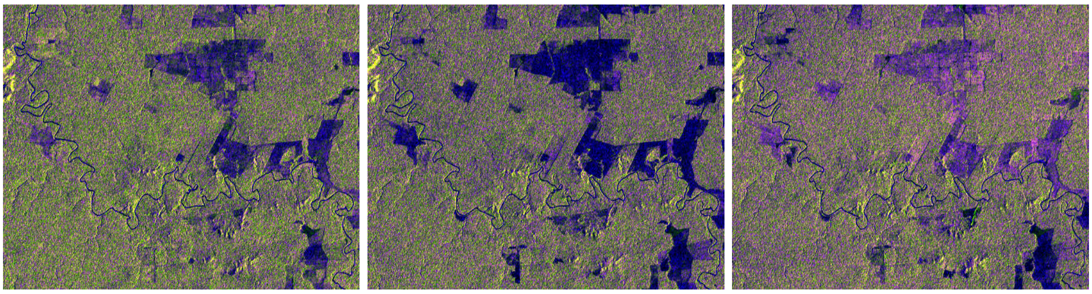
The orientation of these edges allows to conclude the imaging direction. In this case the satellite has been passing to the right of the image as it is a descending orbit. If the exact incidence angle is known, the width of the radar shadow can be used to calculate the tree height from simple trigonometry. Some machine learning algorithms furthermore attempt to use these effects to accurately delineate forest areas by combining images from different viewing angles (Bouvet et al. 2018).



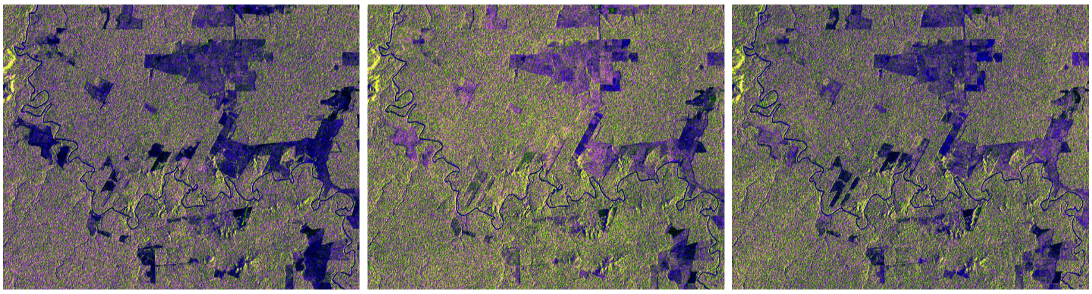
Jan 2017 Feb 2017 Mar 2017



Apr 2017 May 2017 Jun 2017



Jul 2017 Aug 2017 Sep 2017



Oct 2017 Nov 2017 Dec 2017

**Figure 4.1:** This figure shows a subset of the Sentinel-1 data for the study area in Rondônia (red: $C_{VV}$, green: $C_{VH}$, blue: $C_{VV}/C_{VH}$). One image per month is shown out of a total of 30 available images for the year 2017

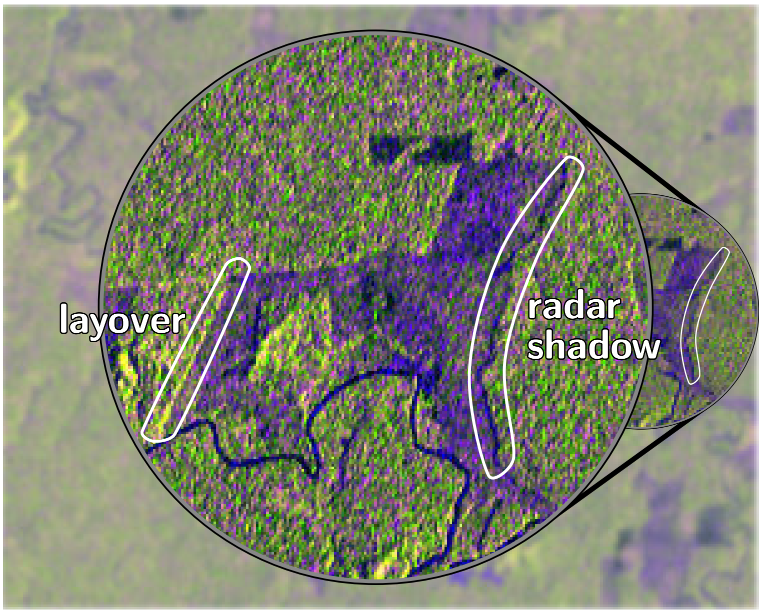
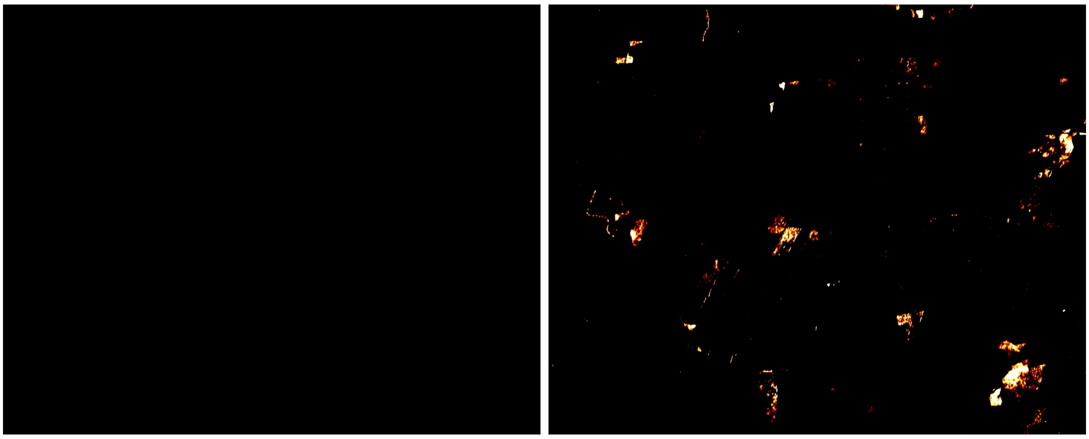


**Figure 4.2:** This close-up of the image from January 2017 reveals areas of layover and radar shadow at forest edges.

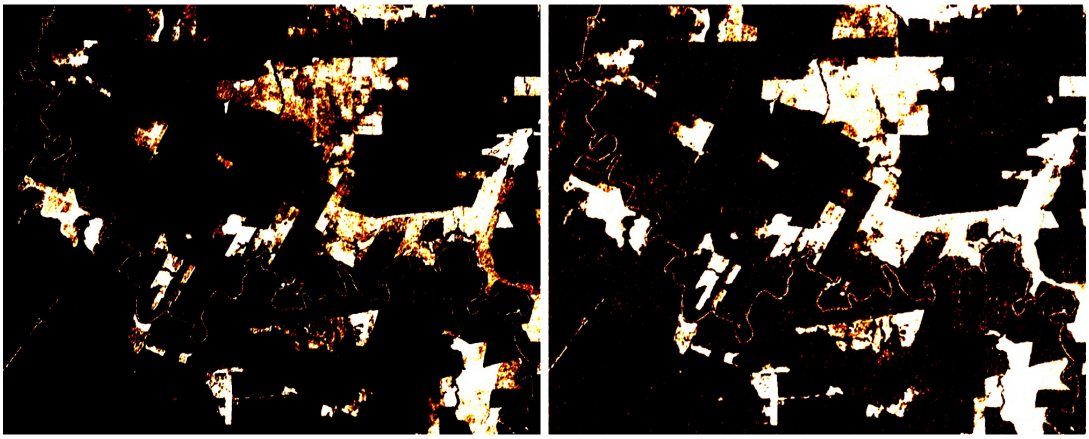
Based on this series of images (a total of 30 time steps), a change detection result according to the algorithm by Conradsen, Nielsen, and Skriver (2016) is presented here. The algorithm was chosen because of its comparatively mathematical derivation and statistical soundness. Furthermore, it takes into account the complete complex-valued and polarimetric properties of each pixel.

Figure 4.3 shows the probability of change as calculated from different ensemble sizes for averaging. The probability is $1 - p(H_0)$ where $H_0 : \mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \ldots = \mathbf{\Sigma}_{30}$. This is equivalent to the first iteration of the algorithm. For pixels that indicate change, the algorithm will then go on to investigate at what time step the change occurred. The ensembles are implemented as simple boxcar windows. Intuitively, a larger ensemble size will average over more pixels, thus remove more noise and should result in a lower probability of change. However, the test statistic itself takes the ensemble size into account (see eq. (4.14)). In other words, the threshold for change detection is adjusted depending on the amount of smoothing. This results in the counterintuitive trend visible in the figure below, where the probability is radically underestimated for small ensemble sizes.



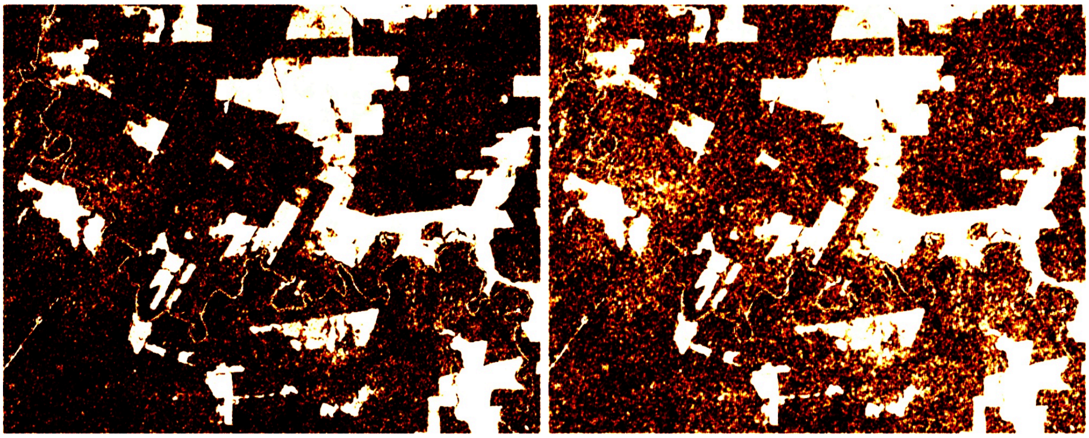
**(a)** $1 - p(H_0)$ for a $1 \times 1$ boxcar window ($n = 1$)

**(b)** $1 - p(H_0)$ for a $3 \times 3$ boxcar window ($n = 9$)



**(c)** $1 - p(H_0)$ for a $5 \times 5$ boxcar window ($n = 25$)

**(d)** $1 - p(H_0)$ for a $7 \times 7$ boxcar window ($n = 49$)



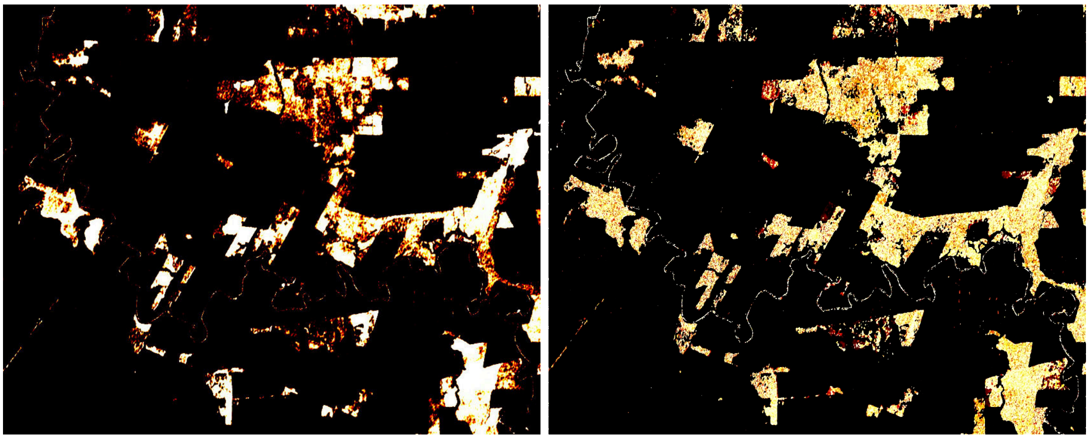
**(e)** $1 - p(H_0)$ for a $9 \times 9$ boxcar window ($n = 81$)

**(f)** $1 - p(H_0)$ for a $11 \times 11$ boxcar window ($n = 121$)

**Figure 4.3:** This figure compares the omnibus change probability $1 - p(H_0)$ for different ensemble average window sizes. The scale ranges from black at 0 to white at 1.

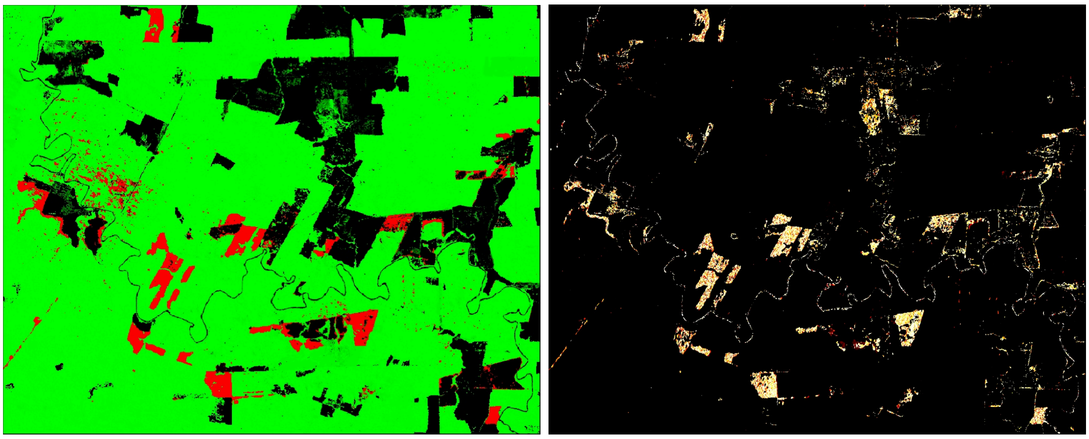
It appears that window sizes anywhere between $5 \times 5$ and $9 \times 9$ could yield reasonable results. Future research is needed to investigate whether change detection results for multiple window sizes can be combined to give an increased confidence of the detected changes.

Further results for the $5 \times 5$ window are summarized in fig. 4.4. **(a)** shows the overall probability of change. Image **(b)** depicts the number of observed changes. Image **(c)** shows reference data for the study area taken from Global Forest Watch (Hansen, Potapov, et al. 2013), and **(d)** is the result from (b) masked by forested area. All images range from 0 (black) to high (white), except for the reference data which show forest cover as green and forest loss as red.



**(a)** Overall probability of change

**(b)** Change frequency, $\alpha = 0.01$



**(c)** Reference data for forest cover and forest loss 2017 (Hansen, Potapov, et al. 2013)

**(d)** Change frequency masked by 2016 forest cover

**Figure 4.4:** These figures show how the result of the omnibus test based change detection algorithm compares to the reference data from GlobalForestWatch.

For the selected area there was an omission error of 39.3% and a commission error of 2.5% compared to the reference data. However, the reference data themselves cannot be assumed to be perfect and will thus have their own errors of commission and omission. Furthermore, it is important to note that the algorithm is a generic change detection algorithm and is not specifically tailored to forest loss. There may therefore be genuine changes present in the image which can be attributed to effects other than forest loss. In general, stable forest is reliably classified as *no change*, according to both visual intuition and the reference data. Agricultural areas have lots of changes, such that it may be possible to use the change frequency to classify between land cover types. Nevertheless, the large omission error suggests that a large fraction of forest loss does not result in a statistically significant change point at this particular configuration of the algorithm.

One way of improving change detection is by taking into account collective trends within a land cover type. Removing these trends should enhance genuine change points. Furthermore, a possible enhancement of the algorithm may employ a better way of defining ensembles than simply using a boxcar filter. By refining the definition of spatial neighborhood in a way that is more likely to group pixels of equal surface cover into the same ensemble we should get more robust change results. This is one of the main focus areas of further research.

### 4.4.1 Change map post-processing (image segmentation)

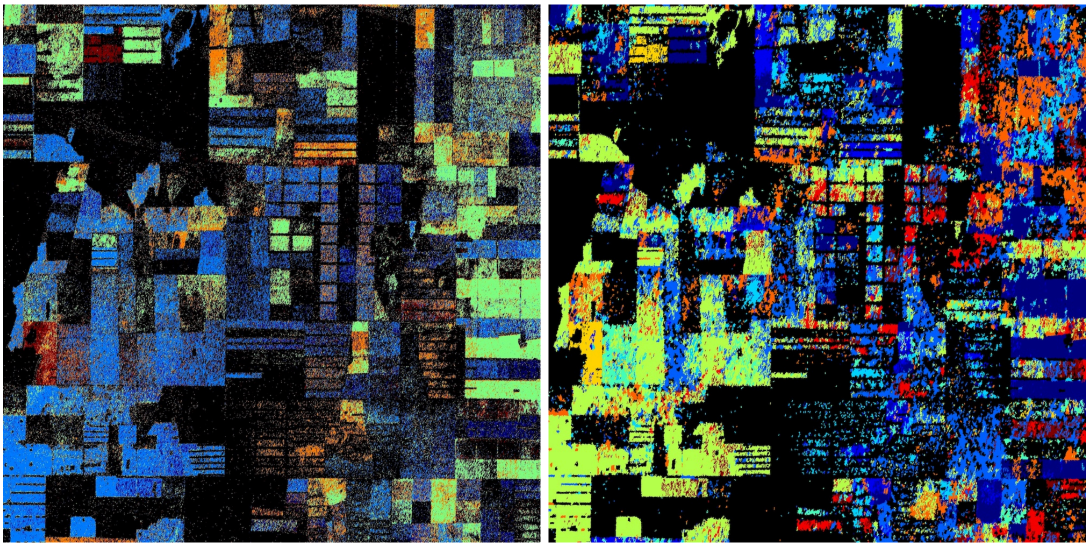
As long as the change detection algorithm does not yield noise-free output maps, it is in some cases necessary to perform some post-processing to refine the output. One of the ways to do this is by using image segmentation to group together areas that show similar change textures but where noise results in a non-homogeneous output. To this end I developed a segmentation algorithm (https://github.com/jnhansen/multiband-segment). This was done in direct cooperation with Ecometrica who required a segmentation method to clean up their change detection results. The methods starts with an initial image segmentation using Felzenszwalb's graph based algorithm (Felzenszwalb et al. 2004) to generate many small candidate segments. The small segments are then iteratively merged using several approaches:

1. A matrix is generated that contains the feature distance between any pair of segments. The feature distance is defined as the Euclidean distance between the average change vector in each segment. A second matrix indicates whether to merge any two segments based on a preset distance threshold. This step only considers adjacent segments for merging, which

is why a second matrix is generated that is `True` wherever two segments share an edge and `False` otherwise. Segments will be merged where both of the latter matrices are `True`.

2. The next step works the same way, but rather than using the Euclidean distance, a t-test checks for equivalence of the change distributions in any pair of segments. Pairs with a $p$-value below a certain threshold are merged.
3. In a third step, KMeans clustering is used to dynamically merge similar segments. Again, adjacency is required.
4. Next, a filter is applied to force-merge small segments of an area below a set threshold with their most similar neighbor.
5. Finally, another KMeans clustering is done. This time adjacency is not required, so this step will reduce the number of segments to a set number.

Each of the steps is optional and available as a parameter from the main call signature. Figure 4.5 shows an example segmentation result for some set of tuning parameters. The colors do not have physical meaning as they are merely segment labels.


**(a)** The input to the segmentation algorithm (shown is only the time of first change)

**(b)** The segmentation output as discrete labels

**Figure 4.5:** This figure shows an example segmentation result for the input on the left.

While the input (fig. 4.5a) arguably looks cleaner and may be all that is required for visual interpretation, it contains a lot of noise, e.g. in the form of pixels with no detected change. This complicates tasks such as computing statistical properties for contiguous image regions. The segmented image enables such computations.

## 4.5 Beyond Pixelwise Change Detection

Change detection is a topic of time series analysis and as such deals with individual time series, i.e. individual pixels in the case where we have time series of image data. However, our data are not only temporal but also spatial and as such may exhibit both spatial and temporal dynamics. The data value of any given pixel needs to be considered within its spatio-temporal context to draw reliable conclusions about the target on the ground. For example, for a fixed target at any given time and location, a pixel value may vary greatly due to environmental conditions such as (soil) moisture, as well as speckle. Speckle can be greatly reduced by averaging over pixels in a spatio-temporal neighborhood that reflects identical or similar targets. Conditions such as soil moisture are likely to affect a whole region rather than individual pixels. Because of this, apparent land cover changes that are due to fluctuations of environmental conditions can be mitigated by considering spatial context.

The question addressed here is therefore the inclusion of spatial context into the change detection process. This could be achieved via direct integration of a spatial component into the change detection algorithm, thus moving away from a pixelwise process. Alternatively, the change detection could be expanded into a two-step process, where the first step would constitute a spatial (or spatio-temporal) filter. That way each pixel already contains information about its neighborhood when a pixelwise algorithm is applied in the second step.

### 4.5.1 Datacubes

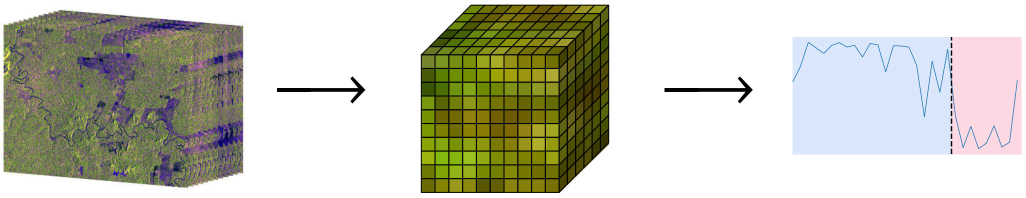


**Figure 4.6:** This figure illustrates how we can rethink a time series of images as a datacube, and how a pixelwise change detection disregards all spatial information.

As a first step towards representing spatial context we can rethink the data structure by moving away from a “time series of images” to a “datacube”, i.e. a multivariate three-dimensional array (two spatial dimensions and one temporal dimension). Datacubes allow capturing the spatiotemporal dependencies within the data. However, change detection remains a pixelwise process, so we lose spatial information as illustrated in fig. 4.6.

### 4.5.2 Spatiotemporal filtering

One way to capture information about the spatial context while maintaining pixelwise change detection is to introduce spatial information into the time series as additional features. The easiest option for generating such features is to apply spatiotemporal filtering. In the general case, a filter recomputes the pixel values $x$ as a function of the values of the pixels in its neighborhood $N(x)$:

$$f\left(x_{ijt}\right) = f\left(\left\{p|p \in N(x_{ijt})\right\}\right) \tag{4.15}$$

For example, a linear filter recomputes $x$ as a weighted average of pixels in its neighborhood:

$$f\left(x_{ijt}\right) = \sum_{p \in N(x_{ijt})} w_p \cdot p \tag{4.16}$$

The simplest example of such a filter would be a boxcar convolution, which is essentially already part of the omnibus test based algorithm constructed by Conradsen, Nielsen, and Skriver (2016) as discussed in chapter 4. This is because the algorithm relies on ensemble averages, which are usually a by-product of multilooking or traditional noise reduction techniques, such as boxcar or Gaussian filters. The aim of such filters is always a noise reduction. They are based on the idea that noise has zero mean and will thus average out over a set of pixels that would otherwise have the same value. Filters differ in how the set of pixels over which to average is determined, as well as in the weights associated with each of those pixels. Boxcar and Gaussian filters are naïve in that they are based exclusively on spatial proximity. While they generally achieve an improved signal-to-noise ratio, this comes at the loss of spatial resolution. More advanced filters thus aim to improve *edge preservation*, where an edge is the genuine boundary between two spatial features rather than a noisy artifact. They generally take into account pixel similarity as well as spatial proximity.

### Non-local means

One such edge-preserving filter is *non-local means* (Buades et al. 2011), which computes filtered pixel values as a weighted average of pixels in the spatial neighborhood, where the weights are determined as a function of color distance. Better measures for pixel similarity could include a combination of spatial proximity and color distance or their temporal correlation. As part of this investigation, non-local means was extended to operate on three-dimensional datacubes. The adapted three-dimensional algorithm is presented in algorithm 1.

One complication when using a generic spatio-temporal filter is the computation of the effective sample size $n_{\text{eff}}$ since all samples are generally weighted. The effective sample size of a weighted sample with weights $w_i$ is given by

$$n_{\text{eff}} = \frac{\left(\sum_i w_i\right)^2}{\sum_i w_i^2}. \tag{4.17}$$

The effective sample size is otherwise known as the Estimated Number of Looks (ENL) (Anfinsen et al. 2009).

**Adapting non-local means for a fixed sample size.** To make non-local means statistically comparable to other filters it is important to allow for a fixed effective sample size for each pixel. However, non-local means effectively yields a sample with some weights $w_i$ that vary from pixel to pixel because the weights are adaptive to the pixel neighborhood. That also means that the effective sample size is different for each pixel. We can think of this as a different *amount* of filtering per pixel. So what if we want to modify the non-local means algorithm such that the effective sample size is constant throughout the image? One way is to adapt the weight assigned to the center pixel itself in such a way that the effective sample size is a given value (this is already reflected in algorithm 1). If we call this weight $w_c$, we get a modified effective sample size

$$n_{\text{mod}} = \frac{\left(w_c + \sum_i w_i\right)^2}{w_c^2 + \sum_i w_i^2}. \tag{4.18}$$

Solving this equation for $w_c$ gives

$$w_c = \frac{1}{n_{\text{mod}} - 1} \left( \sum_i w_i \pm \sqrt{n_{\text{mod}} \left( \left( \sum_i w_i \right)^2 - (n_{\text{mod}} - 1) \sum_i w_i^2 \right)} \right). \tag{4.19}$$

We can thus choose $n_{\text{mod}}$ and calculate $w_c$ accordingly. In this way it is possible to arbitrarily fix the effective sample size for a given pixel to any value $1 < n_{\text{mod}} < n_{\text{eff}} + 1$.

The lower limit follows from

$$\lim_{w_c \to \infty} n_{\text{mod}} = 1 \tag{4.20}$$

and the upper limit follows because the discriminant must be non-negative:

$$\left( \sum_i w_i \right)^2 - (n_{\text{mod}} - 1) \sum_i w_i^2 \geq 0 \tag{4.21}$$

and therefore

$$\begin{aligned} n_{\text{mod}} &\leq \frac{(\sum_i w_i)^2}{\sum_i w_i^2} + 1 \\ n_{\text{mod}} &\leq n_{\text{eff}} + 1 \end{aligned} \tag{4.22}$$

which makes sense intuitively as we cannot increase the effective sample size by more than $1$ if we only add one sample.

Figure 4.7 shows qualitatively how the non-local means filter can drastically improve image quality. The forested areas still exhibit a grainy texture that might be partially due to noise, but is in fact a result of the three-dimensional forest structure and as such remains almost constant over time for stable forest. The improvement of image quality with this filter is more apparent in a video showing the temporal evolution, which cannot be showcased in this document.

**Algorithm 1** Extended Non-Local Means. Originally developed by Buades et al. (2011), I have extended this algorithm to apply to three dimensions and to allow for a fixed effective sample size.

**Precondition:** arr is an $N \times M \times K \times L$ array
**Precondition:** $r$, $f$, $\sigma$, $h$, $n_{\text{eff}}$ are algorithm parameters
**Postcondition:** output is an $N \times M \times K \times L$ array

```
function NLM(arr)
    ▹ Loop over every pixel $i = (i_0, i_1, i_2)$ of the input array
    for $i_0 \leftarrow 1$ to $N$ do
        for $i_1 \leftarrow 1$ to $M$ do
            for $i_2 \leftarrow 1$ to $K$ do
                $w_{\text{sum}} \leftarrow 0$
                $w^2_{\text{sum}} \leftarrow 0$
                $w_{\text{max}} \leftarrow 0$
                for $l \leftarrow 1$ to $L$ do
                    $s_l \leftarrow 0$
                ▹ Loop over every pixel $j = (j_0, j_1, j_2)$ in the neighborhood of $i$
                for $j_0 \leftarrow i_0 - r_0$ to $i_0 + r_0$ do
                    for $j_1 \leftarrow i_1 - r_1$ to $i_1 + r_1$ do
                        for $j_2 \leftarrow i_2 - r_2$ to $i_2 + r_2$ do
                            ▹ Compute the weight between pixels $i$ and $j$, exclude $i = j$.
                            if $i_0 = j_0$ and $i_1 = j_1$ and $i_2 = j_2$ then
                                skip
                            $d^2_{\text{sum}} \leftarrow 0$
                            ▹ Compute the weight between pixels $i$ and $j$ given
                              a window of size $(2f_0 + 1)(2f_1 + 1)(2f_2 + 1)$
                            for $d_0 \leftarrow -f_0$ to $f_0$ do
                                for $d_1 \leftarrow -f_1$ to $f_1$ do
                                    for $d_2 \leftarrow -f_2$ to $f_2$ do
                                        for $l \leftarrow 1$ to $L$ do
                                            $d^2_{\text{sum}} \leftarrow d^2_{\text{sum}} + (\text{arr}[i_0 + d_0, i_1 + d_1, i_2 + d_2, l] -$
                                                        $\text{arr}[j_0 + d_0, j_1 + d_1, j_2 + d_2, l])^2$
                            $d^2_{\text{sum}} \leftarrow d^2_{\text{sum}} / \left[L(2f_0 + 1)(2f_1 + 1)(2f_2 + 1)\right]$
                            $w \leftarrow \exp\left(-\max(d^2_{\text{sum}} - 2\sigma^2, 0)/h^2\right)$
                            $w_{\text{sum}} \leftarrow w_{\text{sum}} + w$
                            $w^2_{\text{sum}} \leftarrow w^2_{\text{sum}} + w^2$
                            if $w > w_{\text{max}}$ then
                                $w_{\text{max}} \leftarrow w$
                            for $l \leftarrow 1$ to $L$ do
                                $s_l \leftarrow s_l + w * \text{arr}[j_0, j_1, j_2, l]$
                ▹ Determine the center pixel weight itself
                if $n_{\text{eff}}$ not given then
                    ▹ Assign maximum weight found to center pixel
                    if $w_{\text{max}} = 0$ then
                        $w_{\text{max}} \leftarrow 1$
                    $w \leftarrow w_{\text{max}}$
                else
                    ▹ Compute the center pixel weight to yield the desired effective sample size
                    $w \leftarrow w_{\text{sum}} + \frac{1}{(n_{\text{eff}}-1)} \sqrt{n_{\text{eff}} \cdot w^2_{\text{sum}} - n^2_{\text{eff}} \cdot w^2_{\text{sum}} + n_{\text{eff}} \cdot w^2_{\text{sum}}}$
                $w_{\text{sum}} \leftarrow w_{\text{sum}} + w$
                for $l \leftarrow 1$ to $L$ do
                    $s_l \leftarrow s_l + w \cdot Im[i_0, i_1, i_2, l]$
                    output$[i_0, i_1, i_2, l] \leftarrow s_l / w_{\text{sum}}$
    return output
```

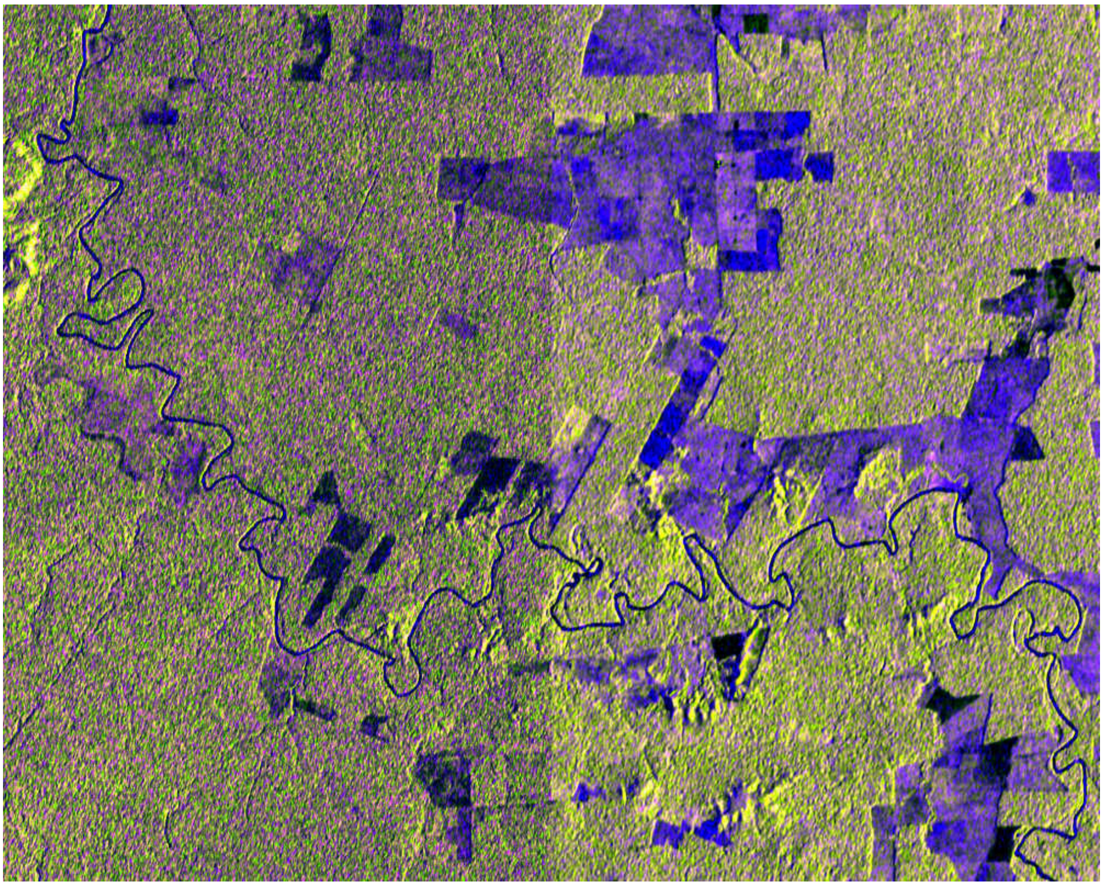


**Figure 4.7:** This figure shows a time slice of the original data (left) compared with the same time slice of the non-local means filtered data (right).

### 4.5.3 Change detection with spatial features

#### Omnibus test

We can compare the change detection results using the Omnibus test (see section 4.4) for a dataset that was filtered with a boxcar filter and a non-local means filter, respectively. For comparability, both filters are applied with an effective sample size of $n_{\text{eff}} = 49$. The results are presented in fig. 4.8.

Clearly, non-local means produces a cleaner looking change map. In particular, the boxcar filtered data result in detected changes along edges, which is easily seen along the river. This is because the boxcar filter smoothes over a window in a way that doesn't preserve edges such that resulting pixel values near the edge are not representative of either the river or surrounding land.

**(a)** Boxcar **(b)** Non-Local Means **(c)** Reference (Hansen, Potapov, et al. 2013)

**Figure 4.8:** The subfigures show the detected changes (white) for boxcar and non-local means filtered datacubes, respectively. The reference shows forest cover (green) and deforestation (red) for the year of interest.

### Bitemporal change detection

While a single spatio-temporal filter may be sufficient, we can in fact apply a series of different filters to capture different aspects of the pixel neighborhood and stack the resulting values to yield a feature vector. This section presents the results from a change detection given such a feature vector. The features were generated by computing the mean and standard deviation of a range of different neighborhoods.

In particular, the features were the mean and standard deviation for spatial window sizes $3 \times 3$ and $5 \times 5$, the mean and standard deviation for window sizes 5 and 7 in the temporal dimension, as well as non-local means filters with the following radii in $(x, y, t)$:

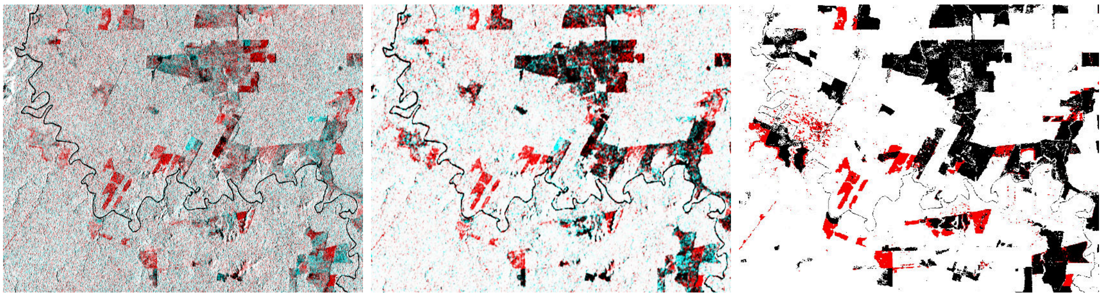
$$(1,1,0), \quad (1,1,1), \quad (2,2,0), \quad (2,2,1), \quad (2,2,2), \quad (3,3,0), \quad (3,3,1), \quad (3,3,2)$$

There is a handy shortcut to compute the windowed standard deviation for each pixel, from the boxcar convolutions of the data as well as the squared data as follows:

$$\sigma_x = \sqrt{\langle x^2 \rangle - \langle x \rangle^2} \tag{4.23}$$

For a bitemporal change detection, we classify each pixel in two different images (at two different times) and detect changes as a change in predicted class. A logistic regression was chosen for the classification and was fitted to some training data where forest was encoded as $1$ and non-forest was encoded as $0$. That way the output of the logistic regression can be interpreted as a forest probability, and the change can be displayed on a continuous scale rather than a binary map.

Figure 4.9 shows the resulting output for the study area at the beginning of 2017 and the beginning of 2018 only. The resulting change maps are thus bitemporal and allow for direct comparison with the reference data.



**(a)** No spatial features **(b)** Spatial features **(c)** Hansen et al. (reference)

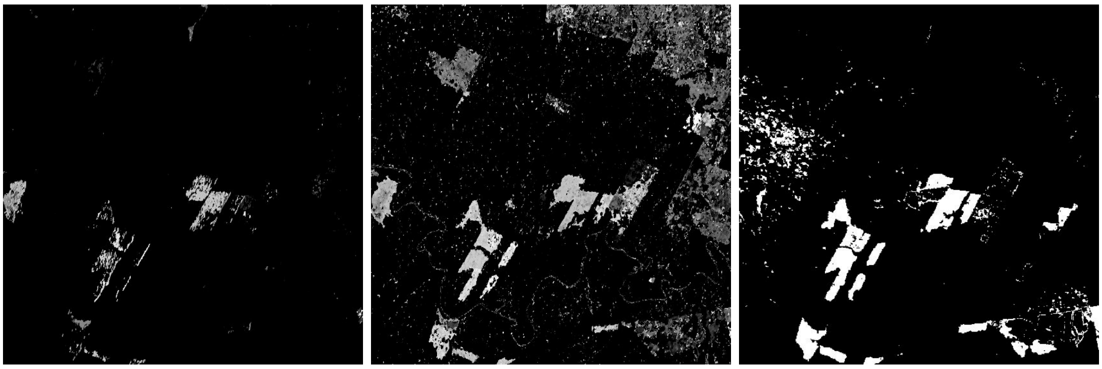
**Figure 4.9:** This figure shows an RGB representation of the detected change for the year 2017, with and without spatial features as well as the reference data. The red channel shows 2017 forest cover, whereas the blue and green channels show the 2018 forest cover.

Rather than showing the difference between forest probability at each time directly, the figure shows the forest probability for each time step in different colo channels. The red and blue/green channels show the forest cover in 2017 and 2018, respectively. That way white areas indicate stable forest, black areas are non-forest, red areas are deforestation, and blue-green areas are forest cover gain. From a visual inspection of the results, fig. 4.9b (i.e. the result including spatial features) comes a lot closer to the reference data. In particular, a significant amount of noise in the forested areas has been removed.

### Time of change from predicted forest probability

Moving on from the bitemporal change detection in the previous section, we can generate the predicted forest probabilities as derived from the logistic regression for the whole time series, and in that way effectively giving a new univariate time series for each pixel. Rather than detecting changes as a transition in predicted class labels, we can then perform a change detection on the generated univariate time series and extract the time of change. This is shown in fig. 4.10 for a subset of the image, with the detected time of change indicated as a grayscale value. The change detection algorithm used was Pruned Exact Linear Time (Killick et al. 2012) (PELT). No reference is available for the exact time of change within the year.

Again, a visual assessment shows an overall better result for the case where spatial features are included in the change detection. The inclusion of these features



**(a)** No spatial features **(b)** Spatial features **(c)** Hansen et al. (reference)

**Figure 4.10:** This figure shows the detected time of change from early (dark) to late (bright), with black indicating no change. For the reference data, the exact time of change is unknown.

appears to increase the change detection sensitivity, reducing the number of omissions but adding a significant amount of false positives. We can conclude that spatial and temporal dependencies are important features of the data. Neither purely pixelwise nor bitemporal change detection approaches make use of the full information captured in the data as they disregard spatial context and temporal characteristics, respectively.

## 4.6 Combining Multiple Data Sources

As technology advances, we constantly see imagery of higher spatial resolution, higher temporal frequency, more sensors, more spectral bands, etc. Combined with the trend towards open data, there is an abundance of remote sensing data available that becomes increasingly difficult to handle. As a consequence, for many applications, the bottleneck has shifted from the lack of data to the development of algorithms that can handle large and disparate data sources.

One keystone in the puzzle of remote sensing algorithms is the ability to combine data from multiple sources, i.e. different sensors or satellites, or even data from space-borne, air-borne, and in-situ instruments. The challenge is not simply one of multivariate data analysis, as these sources typically have starkly different spatial and temporal resolutions. Especially when creating historical records that span the lifetime of multiple missions, the comparability of sensors becomes an issue. Depending on the sensor, slight changes in the setup such as viewing angle, time of day, or spectral bandwidth could mean that the *red* band in one sensor is quite different from the *red* band in another sensor. There have been a multitude of studies looking at optical and radar remote sensing data fusion for land cover mapping (Joshi, Baumann, et al. 2016). In addition, some studies have looked at fusing C-band and L-band radar for forest mapping (Mitchell, Tapley, et al. 2014).

Reiche, Souza, et al. (2013) discuss the fusion of Landsat and ALOS PALSAR data for mapping forest degradation in the tropics. They note that SAR is reliably capable of detecting deforestation and provides higher spatial and temporal coverage than Landsat as it is not affected by cloud cover. Optical sensors, however, allow for better forest degradation assessment through sub-pixel fraction features derived from Spectral Mixture Analysis (SMA) (Adams, Sabol, et al. 1995).

The study area is a site of tropical rain forest in Guyana with an annual cloud cover of $>80\%$. Deforestation in the area is largely driven by agricultural expansion, forestry, and open pit gold mining. The climate is characterized by two annual wet and dry seasons. The ALOS PALSAR data were processed using the Gamma software (Werner et al. 2000), the Landsat data were processed using ImgTools (Souza and Siqueira 2013). Reference data consisted of high resolution imagery from RapidEye, a GIS data set from the Guyana Forestry commission, aerial photographs, and field measurements. 20 samples per forest land cover (FLC) class (forest, degraded forest, non-forest, and water) of $10$ to $25$ pixels each were selected for training, whereas 50 samples per class were selected for validation.

According to the authors, fusion approaches may be categorized into data, feature, and decision level fusion (Pohl et al. 1998), where data level fusion represents the consolidation of different data sources into a single data set before further analysis, feature level fusion represents fusion after data source specific feature extraction, and decision level fusion aims to combine classification results obtained independently from each data source. In this research paper, the authors chose to combine SAR and optical data at the feature level. Specifically, the SAR features were HH and HV backscatter, HH-HV ratio, and dual date difference change index (DCI) (see Equation (4.24)), while the feature extracted from the Landsat data was the Normalized Difference Fraction Index (NDFI) (Souza, Roberts, et al. 2005) that is computed as in Equation (4.25).

$$\mathrm{DCI}_p = \gamma_{p,\mathrm{year1}}[\mathrm{dB}] - \gamma_{p,\mathrm{year2}}[\mathrm{dB}] \tag{4.24}$$

where $p$ is the polarization (HH, HV) and $\gamma = \sigma^0/\cos\theta$ is the radar backscatter parameter ($\theta$ is the incidence angle).

$$\mathrm{NDFI} = \frac{\mathsf{GV}_{\mathsf{shade}} - (\mathsf{NPV} + \mathsf{soil})}{\mathsf{GV}_{\mathsf{shade}} + \mathsf{NPV} + \mathsf{soil}} \tag{4.25}$$

with $\mathsf{GV}_{\mathsf{shade}} = \mathsf{GV}/(100 - \mathsf{shade})$, where Green Vegetation (GV), non-photosynthetic vegetation (NPV), soil, and shade fractions are obtained from an SMA (Adams, Sabol, et al. 1995). For each feature, the class separability was computed as the Jeffreys-Matusita (JM) difference (Laliberte et al. 2012) on the basis of which a decision tree classifier was manually crafted and tuned. The overall accuracy was 86% if water was included in the non-forest class, and 96.6% when degraded forest was considered forest, highlighting the difficulty of detecting degradation compared to full deforestation. 42% of all forest disturbance in the area was due to forest degradation. 80% of all deforestation occurred from undisturbed forest, whereas only 20% occurred from already degraded forest.

Rather than extracting discrete features from the data stream, Reiche, Verbesselt, et al. (2015) fuse bivariate time series (Landsat NDVI and ALOS PALSAR) into a single variable, using their algorithm MulTiFuse. The area surveyed is the *Lololo* forest plantation on Viti Levu, Fiji, and thus lies in the seasonal tropics. The reference data consist of three-monthly harvest records in this plantation from 2000 to 2013, which serve as validation data and provide an initial forest mask. The authors assessed the correlation of Landsat NDVI with the ALOS PALSAR observables (HV, HH, HVHH-ratio) and found the HVHH-ratio to be most strongly correlated to the NDVI data (a correlation coefficient of 0.641 for logged forest), possibly because the ratio may eliminate effects such as soil moisture which heavily influence the SAR backscatter but not the NDVI. Stable forest showed no significant correlation between the NDVI and any PALSAR polarization.

The Landsat processing involved cloud masking with Fmask and atmospheric correction with LEDAPS (Schmidt et al. 2013), whereas the ALOS PALSAR images were processed using the Gamma software package (Werner et al. 2000), including multi-looking, radiometric calibration (Shimada, Isoguchi, et al. 2009), topographic normalization (Hoekman, Vissers, et al. 2010), and geocoding. SAR speckle filtering was done using adaptive multi-temporal SAR filtering (Quegan and Yu 2001).

The MulTiFuse approach consists of the following steps:

1. given time series $X$ and $Y$, linearly interpolate both time series at the times of the other time series to yield $X_{int}$ and $Y_{int}$
2. regression weights are computed for each interpolated point as the inverse difference between the two adjacent values (this is meant to assign small weight to uncertain values)
3. linear regression coefficients for the relationship $Y_{int} = a + bX_{int}$ are determined by least-squares taking into account the regression weights
4. the fused time series $Y_{fus}(t)$ is given by $Y(t)$ where $t \in T_Y$ and by $\hat{Y}(t) = a + bX(t)$ where $t \notin T_Y, t \in T_X$

The authors emphasize that a statistically significant correlation between the two variables is essential for the fusion approach to be sensible. Following the time series fusion, changes were detected using BFAST-monitor (BFM) (Verbesselt, Hyndman, Newnham, et al. 2010; Verbesselt, Hyndman, Zeileis, et al. 2010; Verbesselt, Zeileis, et al. 2012), which models variations of seasonality and trend. However, the authors only modeled the trend as they found little seasonal variability over stable forest. The spatial accuracy and mean time lag was computed as shown in table 4.1.

| | OA | OE | CE | MTL |
|---|---|---|---|---|
| HVHH$_{mt}$ | 92.6% | 13.2% | 2.3% | 4.07 months |
| NDVI$_{MD}$ | 93.1% | | | 1.67 months |
| fused | 95.5% | | | 1.59 months |

**Table 4.1:** This table shows the spatial and temporal accuracy as determined by the authors of the MulTiFuse method.

The effect of increased missing data in the NDVI time series on spatial and temporal accuracy was investigated, from the original 53% up to 95%. While the accuracy of the corrupted NDVI data quickly dropped, the fused time series achieved better accuracy than either of the individual time series in each case.

Building on their previous paper, Reiche, de Bruin, et al. (2015) introduce a Bayesian approach to combining time series from multiple sensors for the purpose of deforestation detection. They use exactly the same data, including the reference data, and pre-processing as before.

In a first step, the authors derived PDFs of forest (F) and non-forest (NF) for each variable/sensor using the available reference data. They tested Gaussian, Gamma, and Weibull models to fit the distributions and assessed the goodness of fit using

Kolmogorov-Smirnov. The best fit was achieved by a Gaussian model in the NF case for both NDVI and HVHH data as well as the F case for HVHH, and by a Weibull model for F when using NDVI. From the forest / non-forest (F/NF) distributions the authors derived the separability as the JM distance (Laliberte et al. 2012).

Next, the authors converted the original time series into a time series of conditional probabilities, where at each pixel and time the conditional probability of non-forest was given by

$$P(\mathrm{NF}\,|s1_t) = \frac{P(s1_t|\,\mathrm{NF})}{P(s1_t|\,\mathrm{NF}) + P(s1_t|\,\mathrm{F})} \tag{4.26}$$

where $s1_t$ is the observation of sensor 1 at time $t$ (the equivalent is true for $s2_t$). The authors denote the resulting time series as

$$s_t^{\mathrm{NF}} = \begin{cases} P(\mathrm{NF}\,|s1_t) & \text{if} \quad t \in T_{s1} \\ \vdots & \quad \vdots \\ P(\mathrm{NF}\,|sn_t) & \text{if} \quad t \in T_{sn} \end{cases} \tag{4.27}$$

At each time step $t$ where $s_t^{\mathrm{NF}} > 0.5$, the probability of deforestation was computed using Bayesian updating:

$$P(D_t|s_{t+i}^{\mathrm{NF}}) = \frac{P(s_{t+i}^{\mathrm{NF}}|D_t)P(D_t|s_{t+i-1}^{\mathrm{NF}})}{P(s_{t+i}^{\mathrm{NF}})} \quad , \quad i = 1, 2, ... \tag{4.28}$$

with $P(D_t|s_t^{\mathrm{NF}}) = s_t^{\mathrm{NF}}$ (this is the case where $i = 0$).

A deforestation event was detected when $P(D_t|s_{t+i}^{\mathrm{NF}}) \geq \chi$ for some threshold $\chi$. It turned out that the optimal $\chi$ depended on the variable: For NDVI, $\chi = 0.975$ yielded the best overall accuracy, while for HVHH, $\chi = 0.5$ was optimal.

The authors determined the following spatial and temporal accuracies when using both the NDVI and HVHH data combined as well as each individual time series ($\mathrm{MTL_F}$ is the mean time lag until the deforestation event is initially flagged, Mean Time Lag (MTL) is the lag until the deforestation event is formally detected):

| | OA | OE | CE | $\mathrm{MTL_F}$ | MTL |
|---|---|---|---|---|---|
| $\mathrm{HVHH}_{mt}$ | 87.3% | 14.5% | 12.2% | 2.2 months | 3.2 months |
| $\mathrm{NDVI}_{MD}$ | 86.7% | 7.8% | 17.8% | -1.2 months | 2.8 months |
| fused | 87.4% | 20.4% | 1.7% | -0.7 months | 1.3 months |

In general, these results are worse than for the MulTiFuse approach. Also, the overall accuracy improved when the amount of missing data in the NDVI time series was artificially increased.

Reiche, Hamunyela, et al. (2018) propose yet another method of fusing SAR and optical data, in particular Sentinel-1, ALOS-2 PALSAR-2, and Landsat NDVI. The authors consider a $10\,000\,\text{km}^2$ area in south-east Bolivia over a two year period between October 2014 and September 2016, i.e. the first two years of Sentinel-1A data. All SAR data were pre-processed using the Gamma software package and were then coregistered to the Landsat NDVI data to an accuracy of 0.5 Pixels. As the research is focused on dry tropical forest, a strong seasonal component is present which is mostly due to changed photosynthetic activity between the wet and dry seasons. The authors removed this seasonal component by spatial normalization in each time step, as proposed by Hamunyela, Verbesselt, et al. (2016).

Change detection was then done for the normalized time series using the Bayesian approach proposed in Reiche, de Bruin, et al. (2015). They used the first year of the two-year period as training data when deriving the PDF for forest (F) and non-forest (NF). The Bayesian updating discussed in Reiche, de Bruin, et al. (2015) was amended for the case when there were multiple variables available for one date, in which case the joint probability was computed.

The authors estimated the spatial accuracy of their detection method by using stratified sampling, while the temporal accuracy, or MTL was calculated as the mean difference between a reference time and the detected deforestation time $t(D_t)$.

Dutrieux et al. (2015) propose a method of incorporating third party data streams into a deforestation detection algorithm. In particular, they look at using coarse resolution NDVI from MODIS data as well as the Standardized Precipitation Index (SPI) from the Tropical Rainfall Measurement Mission (TRMM) to enhance a monitoring system based on NDVI from Landsat. The authors observe that seasonality effects may vary between years and hence render the use of seasonality models inappropriate. In such situations a better approach is to either use spatial context, or include an external regressor capable of representing seasonality. In this case, the lower resolution MODIS NDVI serves as spatial context, whereas the SPI represents a climatic index and therefore may predict seasonality in the NDVI.

For the monitoring the authors used an extension of BFAST (Verbesselt, Hyndman, Newnham, et al. 2010) where the model included terms for external regressors:

$$y_t = \alpha_1 + \alpha_2 + \sum_{j=1}^{k} \sin\left(\frac{2\pi jt}{f} + \delta_j\right) + \alpha_3 x_t + \epsilon_t \tag{4.29}$$

where $x_t$ is the external regressor. Following the BFAST procedure, change is then detected using a moving sum of residuals (MOSUM) between new observations and the model, with a breakpoint declared when the MOSUM exceeds a threshold.

The study area was a site of deciduous tropical dry forest in lowland Bolivia, with clear seasonality and known occurrence of forest loss. A benchmark forest mask was generated from the Landsat based global tree cover data set by Hansen, Potapov, et al. (2013) for the year 2000. All validation was done using visually interpreted high resolution data from Bing maps aided by the TimeSync environment (Cohen et al. 2010). A time lag of six months was found between SPI and NDVI (Vicente-Serrano et al. 2013). Both the SPI and NDVI were interpolated linearly at the dates of the Landsat NDVI time series. While the detection accuracy increased when including the MODIS NDVI, no benefit from including the SPI was found. Possible explanations are thought to be the inadequacy of the SPI for measuring moisture conditions, or the insensitivity of the forests to dry conditions.

## 4.7 Conclusions

In this chapter, I looked in detail at the issue of deforestation detection with SAR data. The use of SAR data was initially motivated by the high cloud cover in forested areas which is hindering the use of optical data.

I reviewed and categorized existing change detection algorithms into model-based and non-parametric algorithms, although other categorizations are certainly possible: the algorithms may be supervised or unsupervised, and they may work on a pair of images (bitemporal) or a full time series (multitemporal).

As an example of a non-parametric change detection algorithm that makes optimal use of the full information content of the dual-polarized SAR data in complex covariance matrix notation, I implemented an omnibus test change detector. The algorithm captured deforestation well, but as expected was not specific to deforestation induced changes alone.

An analysis of the information content in the spatial neighborhood of a pixel revealed improved change detection results when including spatially derived features. I distinguished two basic approaches: (1) the explicit inclusion of additional features representing various properties of the spatial neighborhood; in essence convolutions of the image with different kernels, and (2) the implicit inclusion of the spatial context in the pixel value itself through contrast-preserving spatiotemporal filtering. The second approach was based on the assumption that spatial context is only useful because spatial variability or texture is a noise-like artifact such that spatial averaging is similar to ensemble averaging for a homogeneous land cover. In line with this approach, I extended an existing edge-preserving filter (non-local means) to the three-dimensional case (two spatial and one temporal dimension). I also adapted the filter to allow for a fixed effective sample size – this was crucial to maintain uniform statistical properties throughout the image for use with the change detection algorithm based on a statistical omnibus test. The adapted filter visibly increased both image quality and change detection results.

I also reviewed the literature on data fusion for multi-sensor data, especially a combination of radar and optical. According to Pohl et al. (1998), fusion approaches may be categorized as pixel, feature, and decision level fusion. Pixel level fusion represents fusion at the lowest level, whereas feature level fusion represents data fusion after the extraction of some secondary features such as the NDVI for optical data. In decision level fusion, each data source may be processed independently and a decision function is applied to the information extracted from each source. Most methods reviewed operate at the pixel or feature level and commonly attempt to use radar data to fill in the gaps in optical data where it is missing due to cloud cover.

The following are some of the main challenges in deforestation detection:

(a) Reliable reference data for deforestation and forest degradation with wide spatial coverage, high spatial resolution, and short temporal revisit times are hard to come by.

(b) Significant changes in the backscatter value do not necessarily indicate a change in the land cover or land use. These backscatter variations may instead be due to
    - seasonal variability
    - agricultural growth and harvest cycles
    - variations in soil moisture

(c) The backscatter of forest and non-forest areas can look very similar for some periods during the year, introducing a potential delay in change detection and a risk of introducing false positives if an area is mistakenly classified as forest.

# Chapter 5

# Change from Reference Class Deviation

As discussed in section 1.2, a big challenge for developing and validating algorithms for deforestation detection is the scarcity of reliable reference data. Ideally, the exact time and type of change is known for a large area. However, data of this kind are as of yet infeasible because they require continuous large scale monitoring. For global coverage, the best available forest maps are still only updated at annual intervals (e.g. GlobalForestWatch (Hansen, Potapov, et al. 2013) or JRC Tropical Moist Forests (Vancutsem et al. 2021)).

In order to circumvent the reliance on reference data to train deforestation detection algorithms, one could simply apply change detection algorithms such as demonstrated in section 4.4 that do not rely on training labels. However, this approach has the downside that the detected changes do not necessarily correspond to deforestation. Detected changes may instead reflect land cover transitions other than forest to non-forest, or indeed measurement changes that do not represent a change in the underlying land cover at all, for example:

- Seasonal changes in vegetation
- Growth and harvest cycles in agriculture
- Soil moisture changes due to rainfall

This is because change detection algorithms are in no way specific to any particular type of change but merely pick up statistically significant changes in the raw data, whether or not these changes correspond to a change in the underlying state.

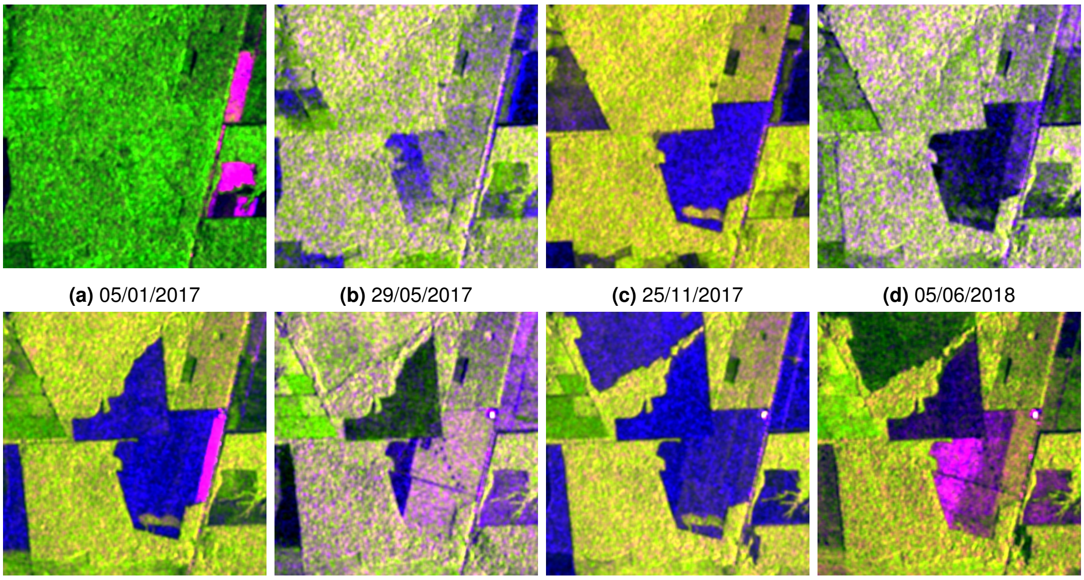
To solve these problems, and to mitigate the disadvantages of both the fully supervised and the fully unsupervised methods, in this section I am exploring a partially supervised approach: Instead of requiring reference data that capture land cover changes and therefore need a temporal component, I have developed a method that only relies on a stationary forest map and thus classifies pixels as either being stable forest or not. Pixels that are not stable forest could include agriculture, urban areas, other forms of vegetation, etc., or pixels that undergo deforestation at some point.

This reference forest map can then be used to detect pixels that deviate from the reference class. This is done by computing a distance metric between the time series of the pixel and the reference forest time series. This can either be computed as the distance to the mean forest time series, or the mean distance to a collection (an ensemble) of individual forest pixels.

The distance can then be computed over time and compared to the expected cumulative distance increase if the pixel was forest. An increase above the expected level suggests a transition from forest to non-forest, with a sudden increase in the distance measure indicating deforestation.

For developing and testing this method in the first instance I chose a small site near Paragominas, Brazil that exhibits very clear deforestation in large patches for the years $2017$, $2018$, and $2019$. Figure 5.1 shows a series of Sentinel-1 images for this study site, in approximately six-month intervals. In addition, fig. 5.2 shows high-resolution optical data for the same area, provided free of charge by Planet through the NICFI program (*Norway's International Climate and Forest Initiative (NICFI)* 2021). The program makes available cloud-free mosaics of high-resolution optical imagery of the tropics every six months from December 2015 to June 2020, and every month since September 2020. The dates shown for the Sentinel-1 were chosen to approximately correspond to the available NICFI images.

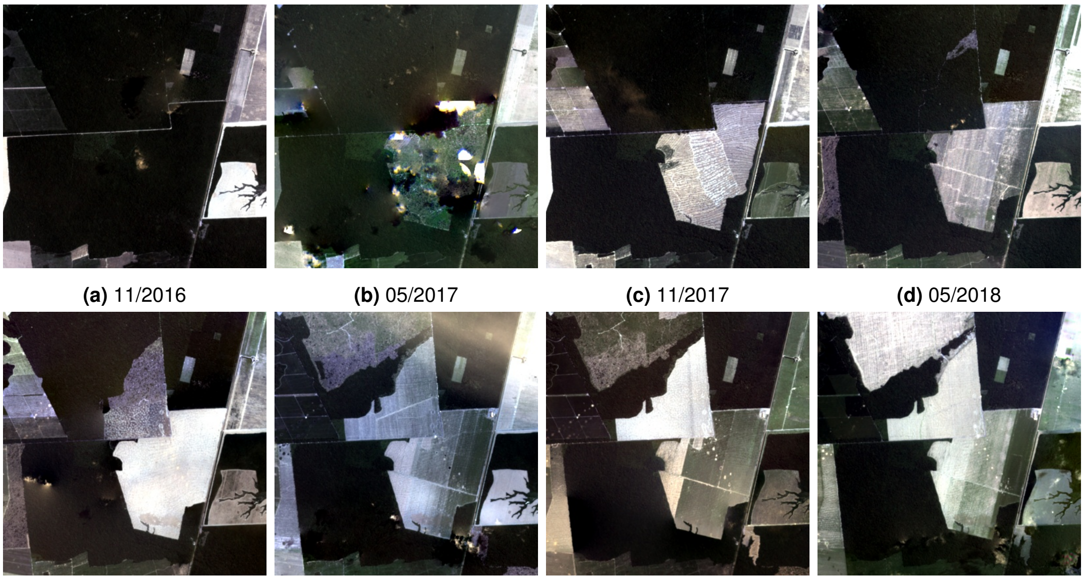
Initially, the NICFI images were planned to be used as validation data for the changes detected with the methods developed in this section. However, the low temporal frequency of six months for the period of time where the changes occur provides little advantage over the widely available annual change maps. In addition, even over six-months periods, the composites were still not entirely cloud free



**(a)** 05/01/2017 **(b)** 29/05/2017 **(c)** 25/11/2017 **(d)** 05/06/2018

**(e)** 02/12/2018 **(f)** 31/05/2019 **(g)** 27/11/2019 **(h)** 25/05/2020

**Figure 5.1:** Sentinel-1 images



**(a)** 11/2016 **(b)** 05/2017 **(c)** 11/2017 **(d)** 05/2018

**(e)** 11/2018 **(f)** 05/2019 **(g)** 11/2019 **(h)** 05/2020

**Figure 5.2:** NICFI images (*Norway's International Climate and Forest Initiative (NICFI)* 2021)

and the compositing potentially hides some of the changes in that period.

The cloud cover problem is further illustrated in fig. 5.3, which shows the entirety of available Landsat imagery over the study site for the years 2017 to 2020. Only a small fraction of the available images are usable due to persistent cloud cover in the area.

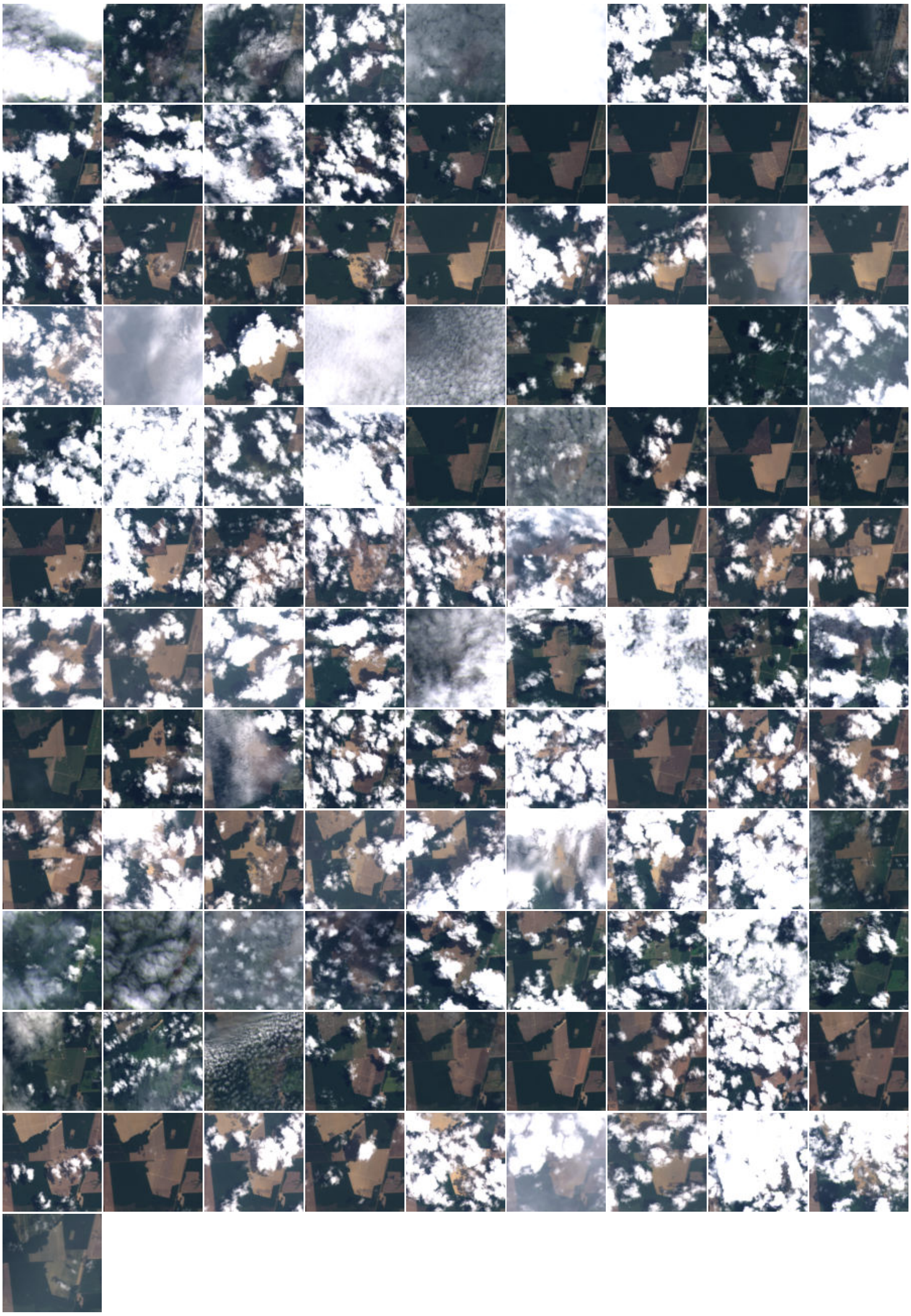


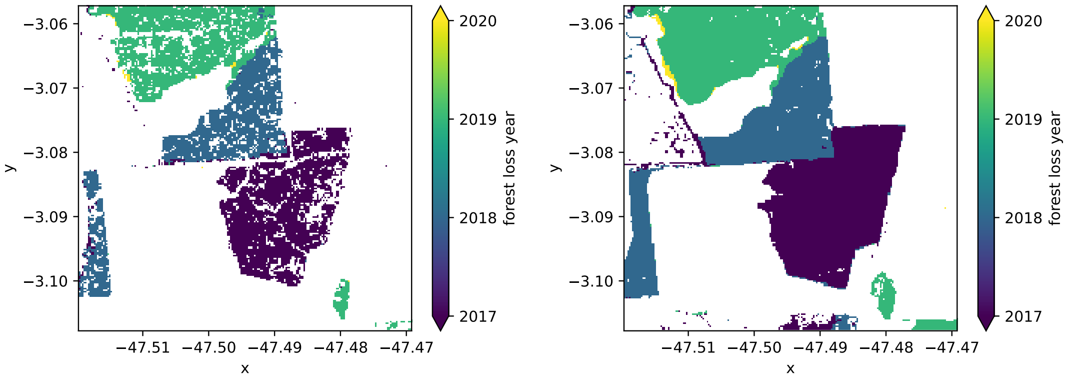
**Figure 5.3:** This figure illustrates the cloud cover problem with optical data. The grid of images shows every available Landsat-8 image from 2017 to 2020 for the study site in Paragominas, Brazil.

## 5.1 Reference Data and Accuracy Assessment

Figure 5.4 shows the annual forest loss events for the study site obtained from two reference data sets:

1. The Tropical Moist Forests product provided by the EC JRC (Vancutsem et al. 2021)
2. The GlobalForestWatch dataset (Hansen, Potapov, et al. 2013)

The change detection results will be evaluated with respect to these maps.



**(a)** JRC reference (Source: EC JRC)

**(b)** Hansen reference

**Figure 5.4:** Reference change maps. The colored pixels indicate the year of forest loss, whereas white pixels indicate no change, which includes stable forest as well as non-forest.

In addition, fig. 5.5 shows a manual segmentation of the study site into polygons that exhibit roughly synchronous behavior throughout the time period. This way, one date of change can be assigned to each area, although some of these areas are deforested in a rainbow pattern, i.e. the change spans a period of time that typically covers a few Sentinel-1 images. The change dates shown in the figure were derived from visual interpretation of Landsat-8 data and represent the latest possible date that the deforestation has occurred. Because of extensive cloud cover in the area (see fig. 5.3), it is possible that the forest loss happened earlier but could not be detected until the date shown.

In general, there are two metrics of interest when assessing the accuracy of any change map that we might term the *spatial* accuracy and the *temporal* accuracy.

The *spatial* accuracy encompasses the precision and recall with which changes are detected in a binary classification setting, i.e. the classification of change vs. no

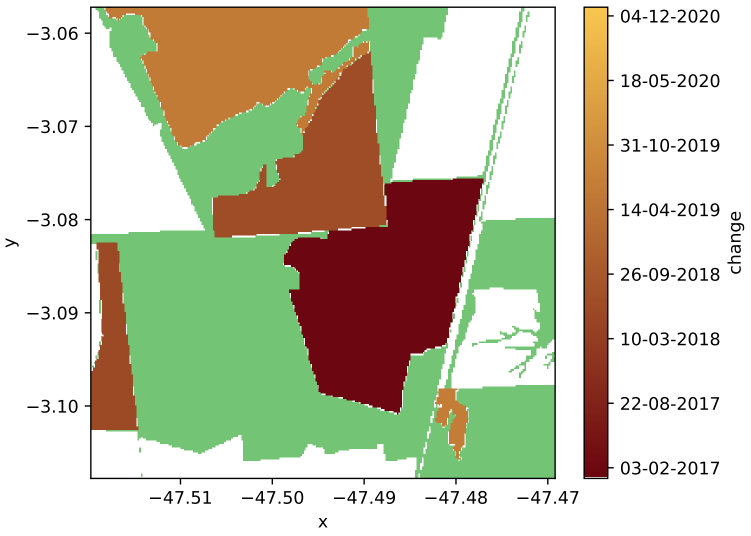


**Figure 5.5:** This figure shows how the study area has been manually segmented into polygons that correspond to areas of approximately equal time of change. The green background represents intact forest for the period of 2017–2020. White areas are non-forest throughout this period, and the areas in shades of brown represent deforestation events at different times in the study period.

change (independent of the time of change). The *temporal* accuracy is a measure of the accuracy of the detected *time* of change. This can be measured as the mean delay of the detected change or the root mean square error (RMSE) of the change time.

While the computation of the spatial accuracy of a change map is straightforward given the reference data shown in fig. 5.4, the temporal accuracy cannot easily be quantified from these validation sources because only the year of change is known. The output of the change detection algorithm will always be a precise time of change that can easily be converted into a year of change. However, while comparing the year of change can give some general insight into the temporal accuracy, it is not a sufficient measure of the uncertainty in the time of change. Not only is comparing the year of change a less accurate measure than comparing e.g. the month of change, it is in fact qualitatively different: if the true year of change for one pixel is 2017 but the detected change is for January 2018, we don't know if this indicates a delay of only one month – which may be an acceptable detection delay – or a whole year, in which case the difference is not merely a delay but likely to be the detection of an entirely different change, i.e. a complete miss of the original first change.

In principle, the resulting map of the year of change could be treated as a multi-class classification problem and the temporal accuracy could be embedded in a multi-class accuracy score such as independent precision and recall rates for each year of change. However, such an approach is not very informative as it treats changes in consecutive years as completely separate classes. This ignores the fact that the distinction between change and no-change has a much higher importance than the distinction between a change occurring in one year vs. another year. Instead, we are only going to compute the mean time lag with respect to the visually interpreted data as a rough indicator of the temporal accuracy, while acknowledging that better reference data are necessary for a true quantification of this uncertainty.

The spatial accuracy is measured as the User's Accuracy (UA) and Producer's Accuracy (PA) as introduced in section 4.2.

## 5.2 Dynamic Time Warping

One popular distance measure between time series is the Dynamic Time Warping (DTW) distance (Sakoe et al. 1978). This algorithm measures the distance between two time series as the amount of *warping* it takes to transform one time series into the other. Because the warping can account for temporal misalignments, Dynamic Time Warping (DTW) is most sensitive to the overall shape of a time series rather than the point-by-point difference in magnitude. This makes DTW more robust and superior to traditional distance metrics in instances where temporal misalignments are common and it is the overall temporal pattern that is of interest. It can further compute distances between time series of different lengths and different temporal sampling. Because of this, DTW is commonly used in fields such as speech recognition (Juang 1984; Koenig et al. 2008; Rabiner et al. 1993; Sakoe et al. 1978) and gesture recognition/motion capture (Choi et al. 2017; Olsen et al. 2018; Riofrio et al. 2017).

DTW compares two time series (temporal sequences)

$$\begin{aligned} X &= x_1, x_2, ..., x_N \quad \text{and} \\ Y &= y_1, y_2, ..., y_M \end{aligned}$$

of lengths $N, M \in \mathbb{N}$ with $x_i, y_i \in \mathbb{R}$.

The notations and definitions presented here are taken from Müller (2007). First, a *local cost* (or *distance*) *measure* $c$ must be defined:

$$c : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}. \tag{5.1}$$

This could be any conventional distance metric, such that $c(x, y)$ is small where $x$ and $y$ are close and large otherwise. Computing the pairwise cost function between all elements of $X$ and $Y$ results in a cost matrix $C \in \mathbb{R}^{N \times M}$ where

$$C_{n,m} = c(x_n, y_m). \tag{5.2}$$

The DTW algorithm then finds an optimal path through this cost matrix with respect to the lowest overall cost. This optimal path is known as the *alignment* or *warping path*. The warping path can be mathematically defined as a sequence $p = (p_1, ..., p_L)$ where $p_i = (n_i, m_i) \in [1 : N] \times [1 : M]$. It must fulfill the following conditions:

1. Boundary condition: $p_1 = (1, 1)$ and $p_L = (N, M)$
2. Step size: $p_{i+1} - p_i \in \{(1, 0), (0, 1), (1, 1)\}$

The second condition also implies monotonicity:

$$n_1 \leq n_2 \leq ... \leq n_L \quad \text{and} \quad m_1 \leq m_2 \leq ... \leq m_L$$

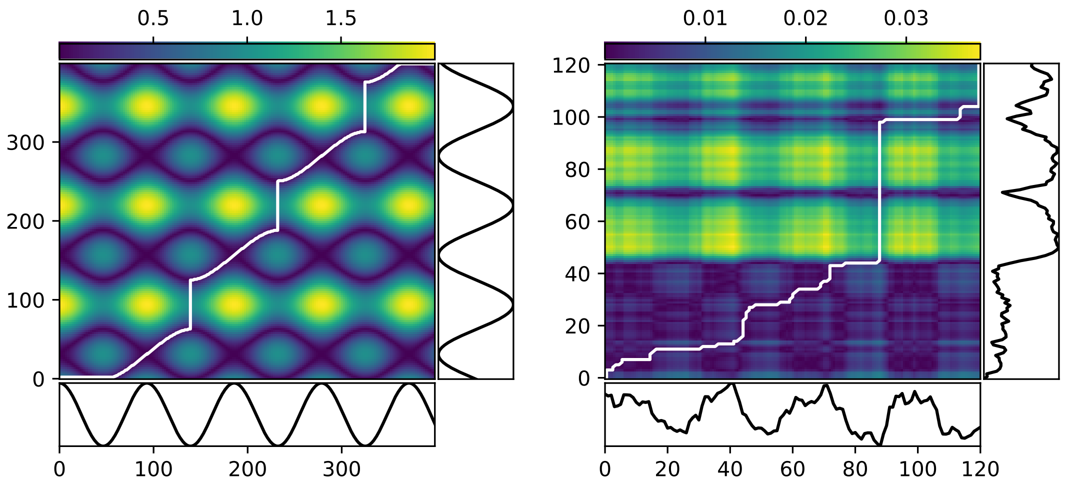
The warping path $p$ defines a mapping between the elements of the two sequences $X$ and $Y$ such that at $p_i = (n_i, m_i)$, the element $x_{(n_i)}$ of $X$ is mapped to the element $y_{(m_i)}$ of $Y$. Because of the step size condition, this can result in a $1$-to-$N$ mapping.

The total cost of the warping path is given by

$$c_p(X, Y) = \sum_{i=1}^{L} c(x_{(n_i)}, y_{(m_i)}). \tag{5.3}$$

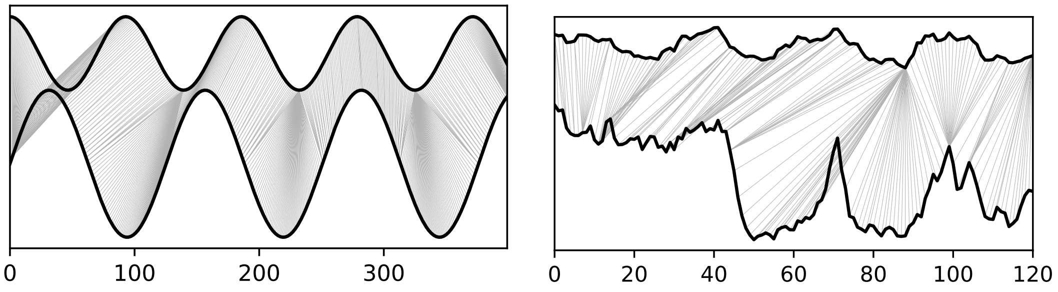
The DTW distance is then defined as the total cost of the *optimal* warping path $p^*$:

$$\begin{aligned} \mathrm{DTW}(X, Y) &:= c_{p^*}(X, Y) \\ &= \min \left\{ c_p(X, Y) \right\} \end{aligned} \tag{5.4}$$



**(a)** Cost matrix A

**(b)** Cost matrix B



**(c)** Alignment pattern A

**(d)** Alignment pattern B

**Figure 5.6:** This figure demonstrates the workings of DTW for two sets of functions. Figure 5.6a shows the cost matrix for two synthetic sinusoidal sequences, and fig. 5.6b shows the cost matrix for two real time series taken from the data set. The optimum warping paths are added in white. Figures 5.6c and 5.6d show the corresponding alignment patterns.

Figure 5.6 demonstrates the cost matrix $C$ (see eq. (5.2)) for two sets of time series, along with the corresponding optimum warping paths $p^*$. It also visualizes the resulting mapping between the two sets of time series as alignment patterns.

Let us now consider some properties of the DTW distance.

First of all, it should be noted that DTW is not technically a distance metric as it does not guarantee the triangle inequality:

$$d(x,z) \leq d(x,y) + d(y,z) \tag{5.5}$$

Furthermore, while equality implies a zero distance

$$X = Y \Rightarrow \mathrm{DTW}(X,Y) = 0 \quad \text{if} \quad c(x,x) = 0, \tag{5.6}$$

the inverse is not always true:

$$\mathrm{DTW}(X, Y) = 0 \nRightarrow X = Y \tag{5.7}$$

It does, however, satisfy the symmetry condition

$$\mathrm{DTW}(X, Y) = \mathrm{DTW}(Y, X) \tag{5.8}$$

as long as the cost function $c(x, y)$ is also symmetric.

Now that we have defined the DTW distance, let us briefly consider an algorithm to obtain the optimal warping path $p^*$ and therefore the DTW distance. One possible algorithm is presented in algorithm 2.

**Algorithm 2** Dynamic Time Warping (DTW) algorithm

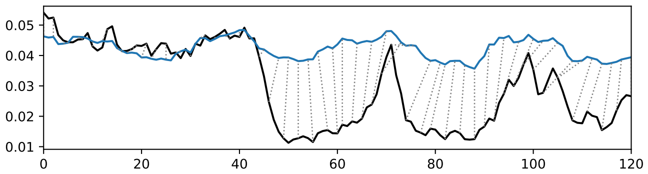
**Precondition:** $X, Y$ are two temporal sequences of lengths $N$ and $M$

```
function DTW($X, Y$)
    $\mathrm{DTW} \leftarrow$ array $[0..n, 0..m]$
    for $i \leftarrow 0$ to $n$ do
        for $j \leftarrow 0$ to $m$ do
            $\mathrm{DTW}[i, j] \leftarrow \infty$
    $\mathrm{DTW}[0, 0] \leftarrow 0$
    for $i \leftarrow 1$ to $n$ do
        for $j \leftarrow 1$ to $m$ do
            cost $\leftarrow d(X[i], Y[j])$
            $\mathrm{DTW}[i, j] \leftarrow$ cost $+ \min(\mathrm{DTW}[i-1, j],$
                                            $\mathrm{DTW}[i, j-1],$
                                            $\mathrm{DTW}[i-1, j-1])$
    return $\mathrm{DTW}[n, m]$
```

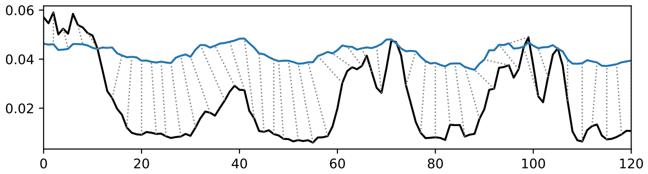
The complexity of the DTW algorithm is $O(NM)$. While small improvements have been achieved (Gold et al. 2018), it has been shown that no strongly subquadratic run time is possible (Bringmann et al. 2015). Some faster implementations exist that give approximate solutions to the DTW problem in linear time, e.g. FastDTW (Salvador et al. 2007).

Figure 5.7 demonstrates the DTW alignment between a forest time series and a time series that includes a change from forest to non-forest.

Figure 5.8 shows the DTW distance between each individual pixel and the refer-



**(a)** Pixel 1



**(b)** Pixel 2

**Figure 5.7:** This figure shows the DTW alignment between a reference forest time series (in blue) and the local time series at two separate pixels (in black). The lines connecting the two time series represent the DTW warping operation.

ence forest time series, obtained as the mean of all pixels that are classified as forest according to the reference data.

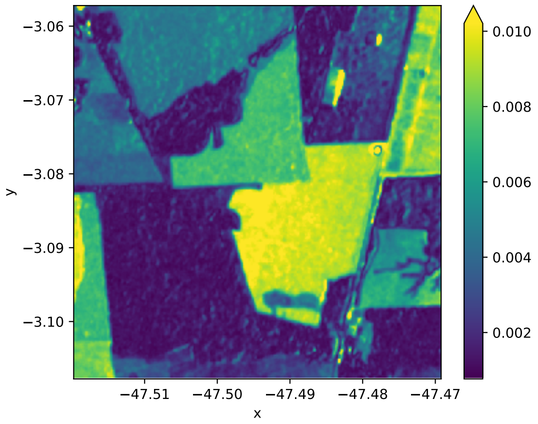
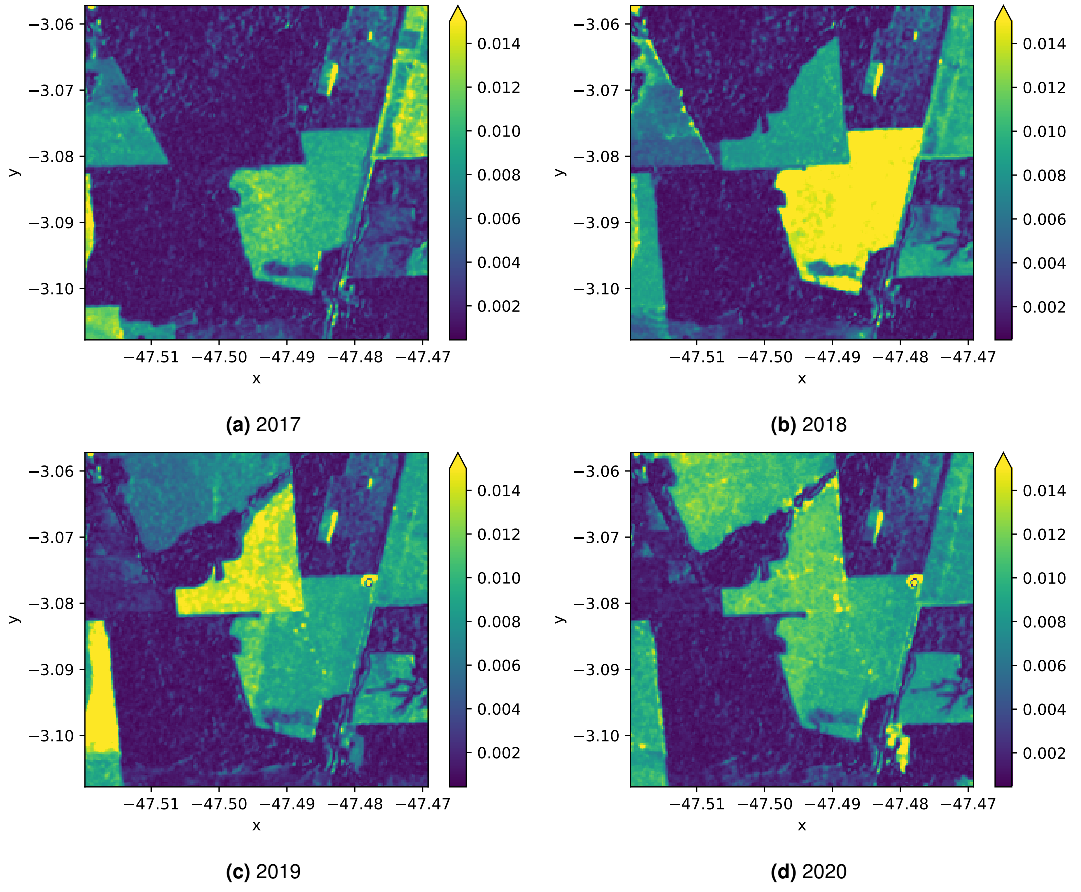


**Figure 5.8:** This figure shows the DTW distance between each pixel and a reference forest time series. Small values (in blue) indicate areas similar to forest, while larger values (in yellow) indicate areas more dissimilar to forest.

While this map gives a good indication as to where stable forest is located, given a small reference set of forest pixels, it does not immediately provide any information about the remaining areas. Pixels with intermediate distance values could be areas of a land cover with temporal backscatter signatures not too different from forest, or they could be areas that were forest for a certain amount of time, and then became non-forest (e.g., through deforestation). To address this ambiguity, and convert the distance map into more of a change detection, one could repeat the distance calculation for a series of subsets of the time series and look at the distance evolution over time. Figure 5.9 shows a series of four such maps for the years $2017$ to $2020$.



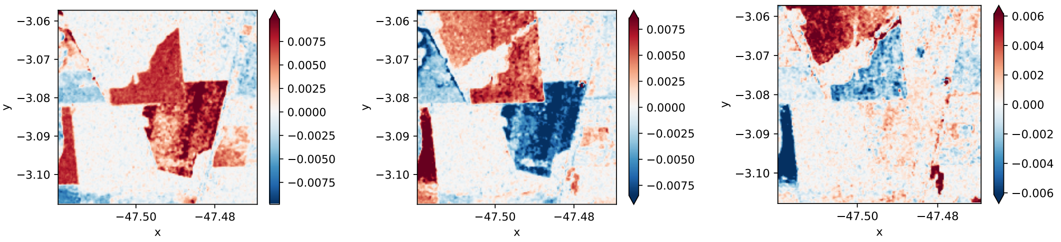
**(a)** 2017 **(b)** 2018 **(c)** 2019 **(d)** 2020

**Figure 5.9:** Each of the subfigures shows the DTW distance between the reference forest time series and the individual pixels for one full year between 2017 and 2020. Blue areas indicate a small distance and therefore stronger similarity to forest, whereas yellow areas represent areas that are more dissimilar to forest.

It now becomes apparent that some of the areas with intermediate values in the previous figure were forest early on (blue areas in fig. 5.9a) and then turned into non-forest later. An example of this is the triangular-shaped area in the upper left quadrant of the image, as well as an elongated area near the lower left edge, that both appear to be forest in 2017 but then change to non-forest at some point in 2018.

Figure 5.10 shows the biannual difference in DTW distance between consecutive years.

Looking at the biannual difference maps, it is clear that there is not sufficient information in these plots alone to detect deforestation: in principle, red areas should indicate a decrease in forest similarity and might thus be declared forest loss, but this does not take into account areas that are not forest to begin with and simply become even more dissimilar to forest in some other way.



**(a)** 2017 – 2018 **(b)** 2018 – 2019 **(c)** 2019 – 2020

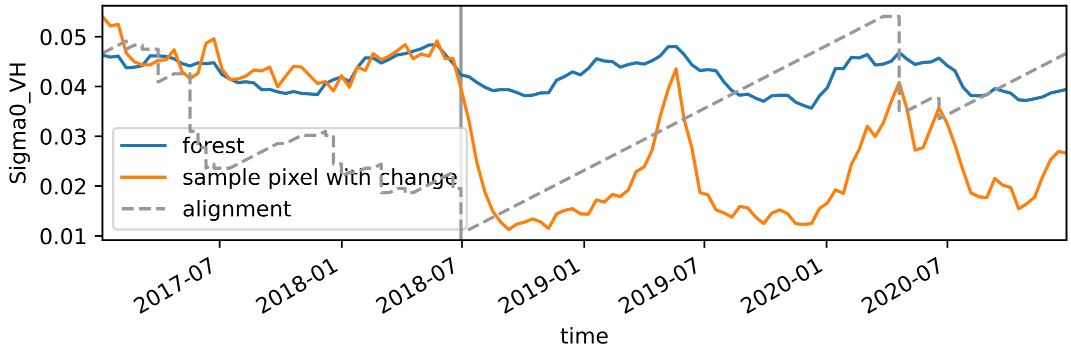
**Figure 5.10:** The three subplots show the biannual difference in forest similarity as measured by the DTW distance (fig. 5.9). Red areas indicate an increase in distance ($\Delta\,\mathrm{DTW} > 0$), while blue areas indicate a decrease in distance ($\Delta\,\mathrm{DTW} < 0$). Areas of forest loss will always be red, but a red area does not inversely imply forest loss. Similarly, blue areas are unlikely to represent forest growth.

While this method can pick up on changes between years, it is unable to pick up on changes near the start or the end of the time series: changes within the first or last year would merely show as areas of an intermediate distance. To mitigate this, one could repeat the same procedure with shorter time series, of course. However, this would ultimately break down as the distance measure becomes less robust for shorter time series. This is because the temporal aspect in itself is what allows us to tell apart forest from non-forest – for any sufficiently short time period, there may be a crop or other land cover type that will look like forest at just that instance. It is the annual behavior that reliably distinguishes different types of land cover, and going to sub-annual time series reduces the overall predictive power.

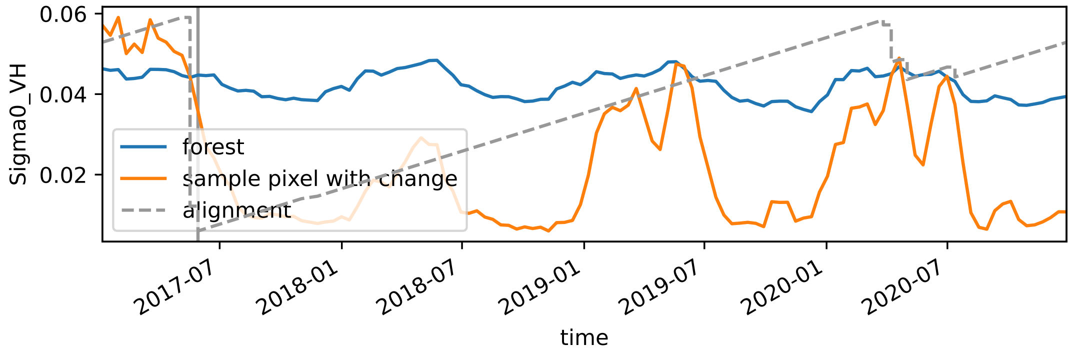
Based on these conclusions, we will now investigate whether a change point can be detected from other aspects of the DTW calculation. Specifically, we will look at the optimal warping path that is a by-product of the DTW distance calculation. As seen in fig. 5.6b, the warping path makes a sharp turn where the area undergoes a change from forest to non-forest.

Figure 5.11 shows two sample time series (in orange) along with the reference forest time series (in blue). The alignment graph (DTW warping path) is shown as a dashed gray line. This corresponds to the white lines in fig. 5.6b, and in fact the first time series is the same as in the example. In this case, however, the warping path is shown rotated by $45°$, corresponding to the relative deviation from the diagonal. Visualizing the alignment in this way is motivated by the fact that the warping path between two identical time series is exactly the diagonal.

By definition, the warping path is the path through the cost matrix with minimum total cost. We can think of the cost as the *time* it takes to cross each element of



**(a)** Detected change (pixel 1)

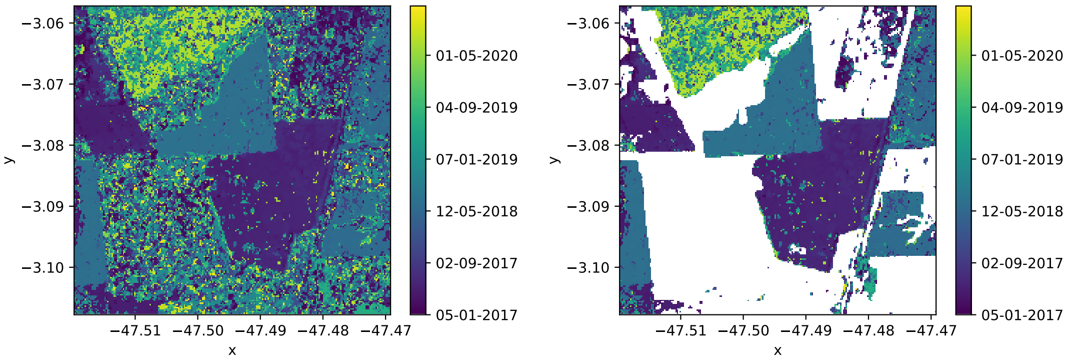


**(b)** Detected change (pixel 2)

**Figure 5.11:** This figure illustrates how the change point can be read off the DTW alignment graph (dashed gray line).

the cost matrix such that the resulting warping path is the *fastest* path through the matrix. This is analogous to the path that light takes through a series of media with different optical density, also known as refraction. This analogy is helpful for considering what happens at a boundary between two regions of low and high cost. For a ray of light, this would be at the interface between two optical media, and the resulting path exhibits a sharp corner that can be precisely calculated. The result is a path that maximizes distance travelled in the less optically dense (faster) medium and minimizes the distance travelled in the more optically dense medium. Equivalently, if the cost matrix has one block region of low cost and another block region of high cost, the warping path will be such that a large part of the path is spent in the area of low cost and a small part in the area of high cost, resulting in a sharp corner just like in light refraction. The reason we consider a cost matrix that is made up of two blocks with different cost is that this scenario exactly represents the type of change we are interested in: if the pixel exhibits forest loss, the distance between the time series and the forest time series will be small initially, and large after the change event. Of course, the regions of low and high cost are not as uniform as in the refraction example, such that the resulting path through either region is not just a straight line – but the general principle still applies (as demonstrated in fig. 5.6b). In particular, for two mostly uniform cost regions, the corner at the interface between both regions will generally be the point that is furthest from the diagonal, and can thus be calculated as the minimum of the rotated warping path (marked as a vertical gray line in fig. 5.11). The figure demonstrates qualitatively for the two sample time series that this method is indeed able to detect the change point.

Figure 5.12 shows the result when this method is applied to each individual pixel of our study region. The left subfigure (fig. 5.12a) shows the original change map, whereas in the right subfigure (fig. 5.12b) a minimum distance threshold has been applied to remove pixels that are likely stable forest.

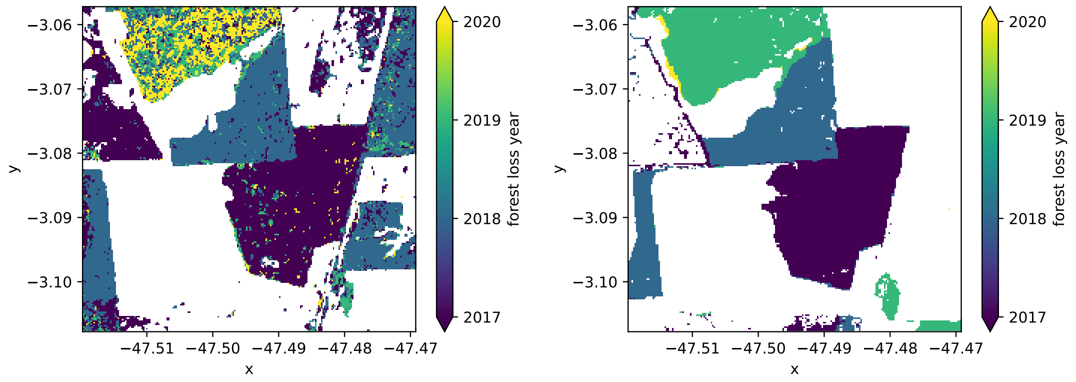


**(a)** Original change map

**(b)** Change map with minimum distance threshold

**Figure 5.12:** This figure shows a map of change points as detected from the corner in a DTW warping path. Figure 5.12a shows the detected change point for each pixel, whereas fig. 5.12b excludes all pixels that do not meet a minimum distance threshold and are thus likely to be stable forest.

Overall, the resulting change map looks promising as it picks up on most changes and appears to do a particularly good job at minimizing false positives in forested areas. However, a direct comparison of the predicted year of change with a reference change map (fig. 5.13) reveals that this method not only picks up areas of forest loss, but also areas that were never forest to begin with. This is because a change point is initially declared for every pixel. A mask is applied only later as a minimum distance threshold to eliminate pixels that are stable forest as indicated by a small overall distance. This does not remove non-forest pixels, which exhibit a large overall distance – though not necessarily larger than pixels with change points, depending on the time of change and type of land cover, so they cannot be removed by distance thresholding. This results in a PA of 87.6% and a rather low UA of 63.6%. In addition, the areas of detected change are rather noisy and much of the change in year 2019 is mistakenly classified as having changed only in 2020.



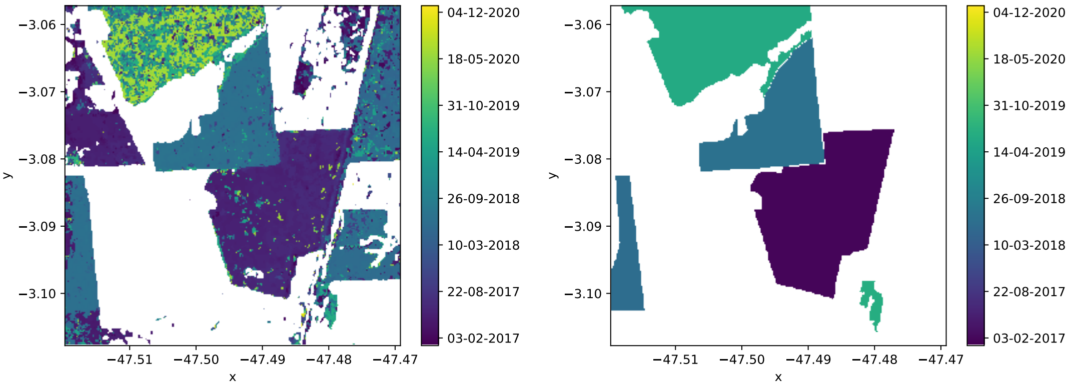
**(a)** Year of change from DTW alignment

**(b)** Reference change map

**Figure 5.13:** This figure shows the detected year of change alongside a reference change map. The colored pixels indicate the year of forest loss, whereas white pixels indicate no change.

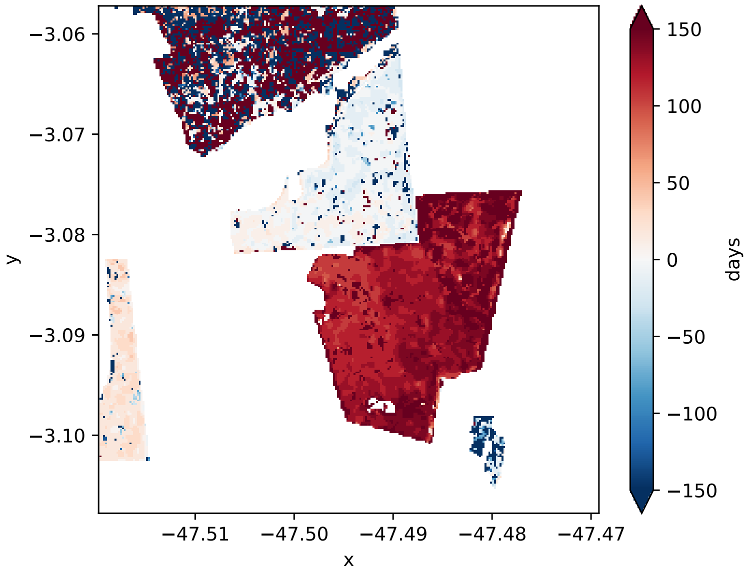
Figure 5.14 shows a comparison of the detected date of change and the reference date of change from the visually interpreted change map, including a map of the corresponding difference in days for each pixel. In general, areas where the change happened near the beginning or the end of the time series appear to have much larger temporal errors than areas with changes in the middle of the time series. However, the number of different areas and amount of data in this study site is not sufficient to draw any reliable conclusions in this regard.

Ultimately, while there is some value in this approach, it is questionable whether DTW is in fact the best metric to quantify forest similarity. Many classic applications of DTW are in areas where e.g. synchronicity of two time series is not given. In this case, however, forest pixels are expected to behave in roughly the same way at each time step, without any relative lag or other temporal warping. An exception would be the comparison of two forest pixels in different areas of the world, where the difference in seasonality may amount to some transformation akin to time warping. It was expected that the characterization of time series similarity by the DTW measure would prove to be more robust than conventional distance metrics given the temporal aspect of the data. However, it appears that for our analysis, DTW is unlikely to provide much benefit over other metrics that compare time series point-by-point. Other metrics will be evaluated in the following section.



**(a)** Date of change from DTW alignment

**(b)** Reference change date



**(c)** Change detection lag in days

**Figure 5.14:** This figure shows the detected date of change alongside a reference change map. Figure 5.14c shows the detection delay with respect to the reference map. Red areas indicate a positive delay, i.e. the change was detected too late, while blue areas indicate a negative delay, i.e. the change was detected too early.

## 5.3 Cumulative Distance Measures

While the results from the DTW warping path show some promise, the advantage of using the DTW distance over more traditional distance metrics is questionable. To address this, we now want to explore how the DTW results stack up compared to changes detected from simpler distance metrics. While the DTW algorithm directly provides a temporal component that the point of change can be derived from (the warping path), computing e.g. the Euclidean distance between two time series provides a scalar without a temporal aspect. We therefore need to reintroduce a time series aspect to the distance calculation in some other way that reflects the distance evolution over time. The underlying assumption is that for pixels that undergo a change from forest to non-forest, the distance between this time series and a reference forest time series will be low initially, and increase post-change.

One way to measure the distance evolving over time is to compute the cumulative distance for each time step. For two time series

$$X = x_1, x_2, ..., x_N \quad \text{and}$$
$$Y = y_1, y_2, ..., y_N$$

(following the DTW notation, but assuming equal length here), let us denote the first $t$ elements of the time series as

$$X_{[1,t]} := x_1, x_2, ..., x_t \quad \text{and}$$
$$Y_{[1,t]} := y_1, y_2, ..., y_t$$

for $t \leq N$. The cumulative distance $D(X, Y) : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}^N$ is then given by

$$D(X, Y) := d\left(X_{[1,1]}, Y_{[1,1]}\right), d\left(X_{[1,2]}, Y_{[1,2]}\right), ..., d\left(X_{[1,N]}, Y_{[1,N]}\right) \tag{5.9}$$

for some distance metric $d(X, Y) : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$.

In particular, we will consider the following distance metrics:

- the Euclidean distance

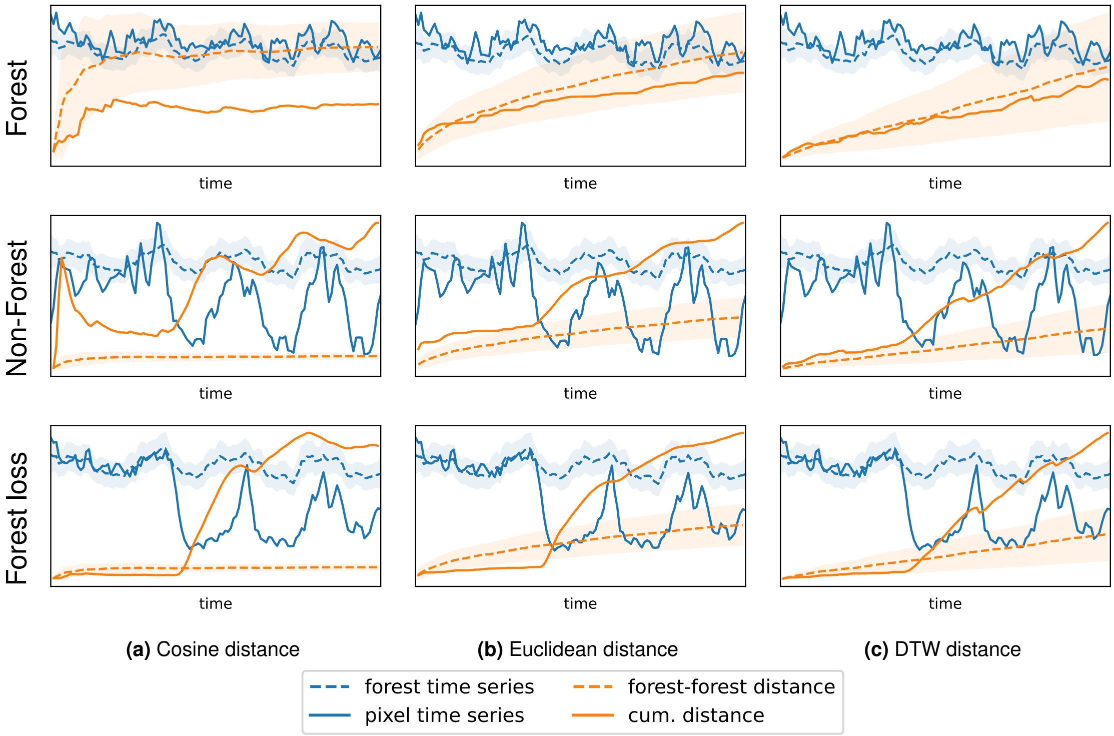
$$d(X, Y) = \sqrt{\sum_{i=1}^{N} (x_i - y_i)^2}, \tag{5.10}$$

- the cosine distance

$$d_{\text{cos}}(X, Y) = \frac{X \cdot Y}{||X||\,||Y||} = \frac{\sum_{i=1}^{N} x_i y_i}{\sqrt{\sum_{i=1}^{N} x_i^2}\sqrt{\sum_{i=1}^{N} y_i^2}}, \tag{5.11}$$

- and the DTW distance (eq. (5.4)) for a direct comparison.

Figure 5.15 shows three sample time series from the study site (as a solid blue line) along with the reference forest time series (dashed blue line), as well as the cumulative distance between the pixel and the forest time series over time (solid orange line). Also shown is the mean cumulative distance between two forest pixels, which is the expected cumulative distance if the pixel were forest. This is shown as a dashed orange line along with the standard deviation as a shaded orange area.



**(a)** Cosine distance **(b)** Euclidean distance **(c)** DTW distance

**Figure 5.15:** This figure demonstrates the behavior of the cumulative distance for three different distance measures: cosine (fig. 5.15a), euclidean (fig. 5.15b), and DTW (fig. 5.15c). The blue lines show the pixel time series (solid line) and a forest reference (dashed line), while the cumulative distance is shown in orange. The dashed orange line is the expected cumulative distance if the pixel were forest.

As expected, the cumulative distance of the forest pixel (first row) stays within the shaded orange area, indicating a high confidence that this is indeed a forest pixel. The expected behavior of the cumulative distance varies between the different distance metrics. For the cosine distance, it levels off and plateaus quickly – this is because the angle between two feature vectors both representing forest will reach an equilibrium and not increase with additional observations. The Euclidean distance accumulates with the square root of the number of observations, and the DTW distance increases linearly.

For the non-forest time series (second row), the difference between the three distance measures is most apparent. In all three cases, the cumulative distance eventually ends up outside of the area representing $\pm 1\sigma$ of the forest ensemble, but for the Euclidean distance and especially the DTW distance, the time series are considered very similar to forest for an initial period. Looking at the time series, this pixel does indeed look very similar to forest for the first period of time, until it starts exhibiting more drastically different behavior. However, while the mean of the time series is similar to forest, it shows higher fluctuations that would be expected. The cosine distance is able to pick up on this more readily.

Finally, for the pixel undergoing forest loss, all three distance measures show a cumulative distance well inside the expected bounds while there is still forest at this pixel, and a clear increase once the pixel has changed. However, there does appear to be a short lag between the point where one would intuitively place the change point and the corner in the cumulative distance curve, for all three distance measures.

## 5.4 Change Detection from Cumulative Distance

Now that we have introduced different distance measures and the rationale for looking at the cumulative distance function, we need to develop a method for detecting a change in this curve. In this section, we are exploring a range of such methods and will be assessing the advantages and disadvantages of each. In general, all of the methods take the difference between the cumulative distance of the pixel and the expected forest-forest distance as input. For convenience, this difference in distance will be referred to as the *excess distance* in the following paragraphs. The methods are presented in the order of sophistication.

**Crossing** With this method, the point of change is declared to be the first point in the sequence where the cumulative distance exceeds the expected value if the pixel were forest. The expected value can be the mean forest-forest distance or, for a more confident result, any higher percentile of the forest-forest distance.

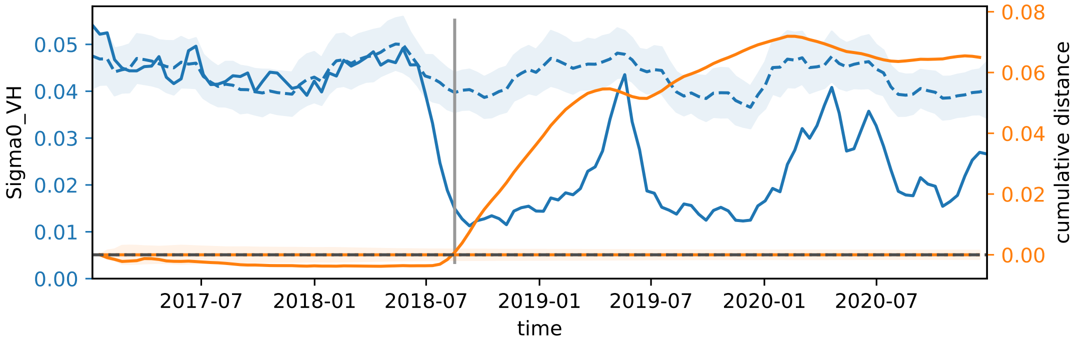


**Figure 5.16:** Illustration of change point detection by finding the crossing point of the pixel-forest distance with the expected forest-forest distance. The distances shown are the excess distance above the expected distance and so the change point is the zero-crossing of the cumulative distance. In this example, the cosine distance metric is shown.

**Argmin** This method is similar to the crossing point method, except the change point is placed at the minimum rather than the zero-crossing of the excess distance. The change point is therefore simply the point where the cumulative distance undercuts the expected forest-forest distance by the maximum amount.

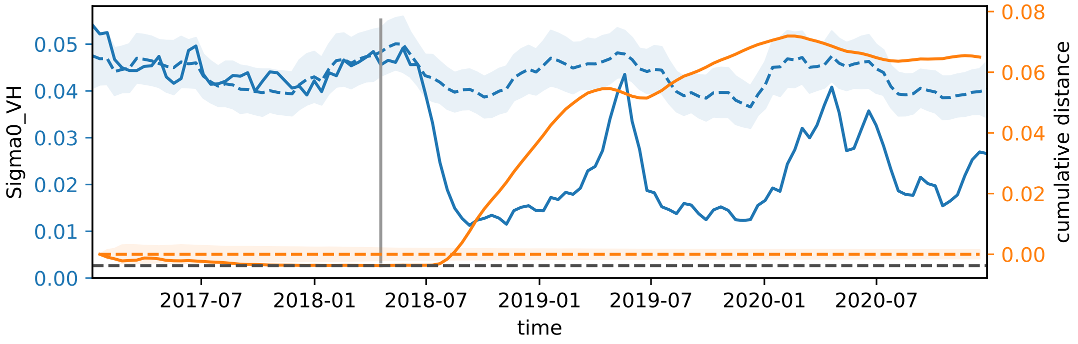
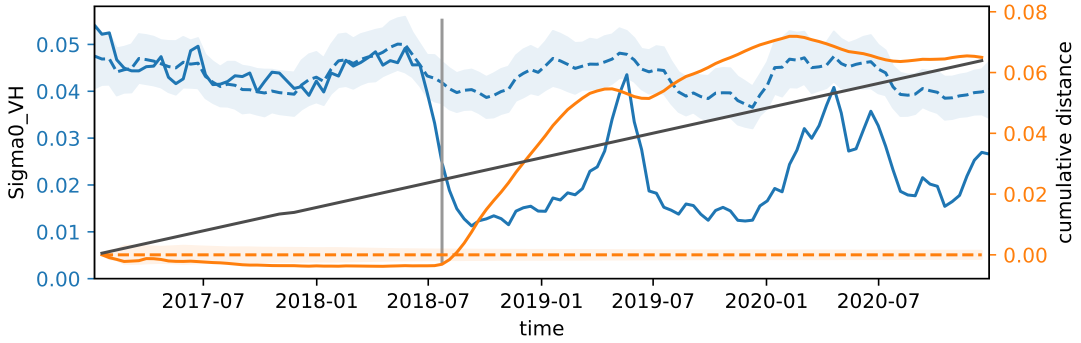


**Figure 5.17:** This figure illustrates the argmin method. The change point is placed at the time step where the excess difference is lowest. The minimum value is indicated by the dashed gray line.

**Linear fit** This method is similar to the DTW warping approach described in the previous section. The cumulative distance curve is realigned with respect to a line representing a linear increase from zero at $x_1$ to the last value of the cumulative distance at $x_N$. The change point is then placed at the point where the excess distance is lowest with respect to this linear fit.

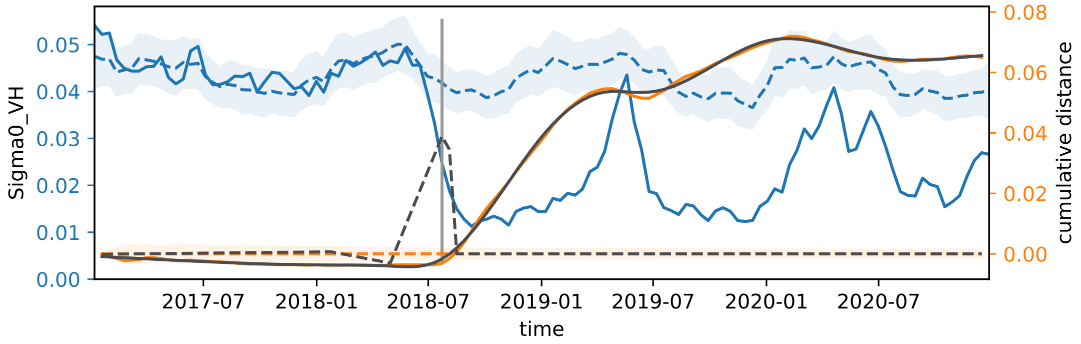
$$x_{\text{change}} = \operatorname*{argmin}_{x_i} d\left(X_{[1,i]}, Y_{[1,i]}\right) - d\left(X_{[1,N]}, Y_{[1,N]}\right) \frac{x_i}{x_N} \tag{5.12}$$



**(a)** cosine

**Figure 5.18:** This figure demonstrates the linear fit method. The fit is shown as a black line and the change point is marked where the cumulative distance is minimum with respect to this line.

**Splines** Another approach is to fit splines to the excess distance. The main advantage over using the raw time series is that it turns the data into a smooth function, allowing us to robustly compute derivatives. In particular, we can detect the point of change as the point of maximum curvature (largest second derivative). This requires appropriate smoothing in the spline fit to avoid large curvature values in smaller fluctuations.



**(a)** cosine

**Figure 5.19:** This figure shows the workings of the splines method. The splines fit is shown as a solid black line, and the second derivative, i.e. curvature, as a dashed black line. The change point is placed at the point of maximum curvature.

**Curve fit** We now explore a more elaborate method for detecting the point of change by fitting a curve to the difference in the cumulative distance. In particular, we are fitting the following S-shaped curve to the cosine distance

$$f(x) = a + b \cdot \tanh(c(x - d)) \tag{5.13}$$

with the four free parameters $a, b, c, d$. The curve fit then allows us to find a change point by setting a threshold increase in the cumulative distance. The S-curve transitions from a lower level $(a)$ to a higher level $(a + b)$, and we might declare the change at e.g. a 10% increase, i.e. the point where $f(x) = a + 0.1 \cdot b$. The exact threshold can be tuned to minimize the delay in the change detection.

For the euclidean distance metric, we are using a different fit that better models the behavior of this metric by combining a piecewise linear function with a square root.

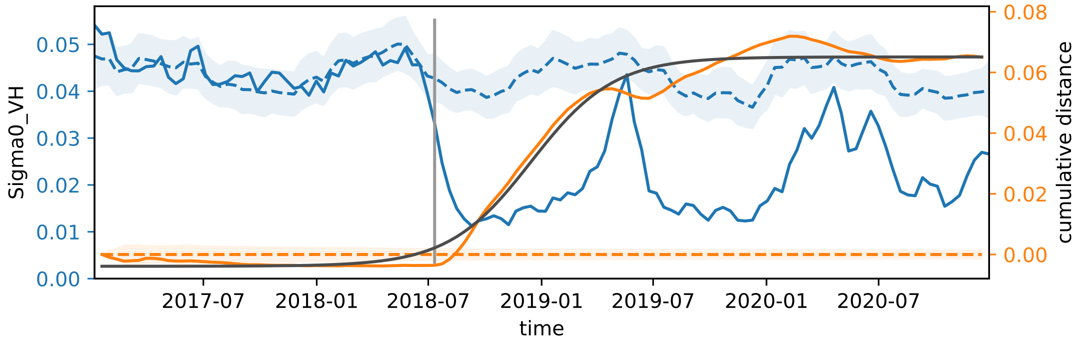
$$f(x) = \begin{cases} k_1(x - x_0) + y_0 & \text{if} \quad x \leq x_0 \\ k_2\sqrt{x - x_0} + y_0 & \text{if} \quad x > x_0 \end{cases} \tag{5.14}$$

with the free parameters $x_0, y_0, k_1, k_2$. In this case, the change point can be easily read off as the parameter $x_0$.

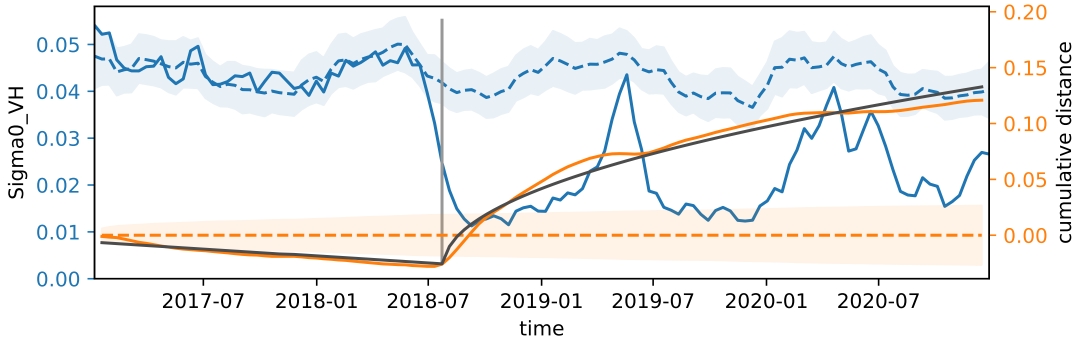
Finally, figs. 5.21 and 5.22 show a direct comparison of the change points detected by the various methods for four different pixels, using the cosine and euclidean distance.

Figures 5.23 and 5.24 show the resulting change maps for every combination of distance metric and change point detection.

The change maps reveal some of the advantages and disadvantages for each of the considered distance metrics and methods. One observation is that all methods detect a change in the area near the upper left edge of the image and also near the upper right edge, where neither of the reference data sets lists these areas as undergoing a change in the time period from 2017 to 2020. Looking at the time series for these areas, at least one of them does indeed look indistinguishable from forest in the beginning of the time period. Only in the optical image in fig. 5.2a do we see that the area has already been deforested at this point. This could indicate an inherent delay in the change detection, and we would need to look at earlier images to see if the detected change would in fact date back to an earlier year in that case. This is also an indication that some minimum amount of time is required to declare a change with certainty.



**(a)** curve fit method with a cosine distance

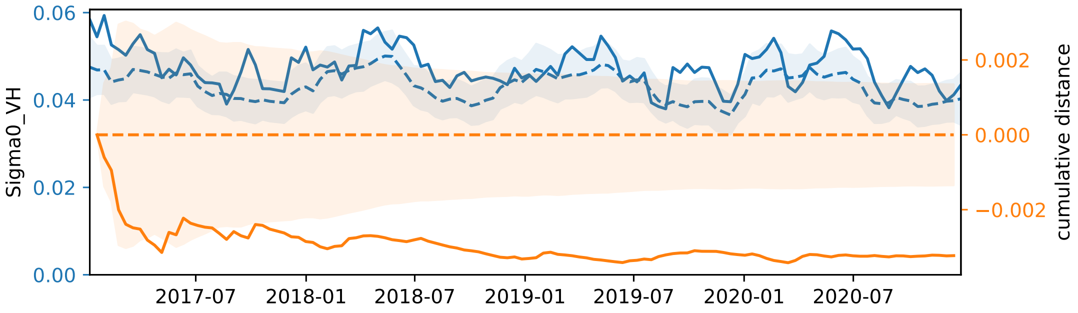


**(b)** curve fit method with a euclidean distance

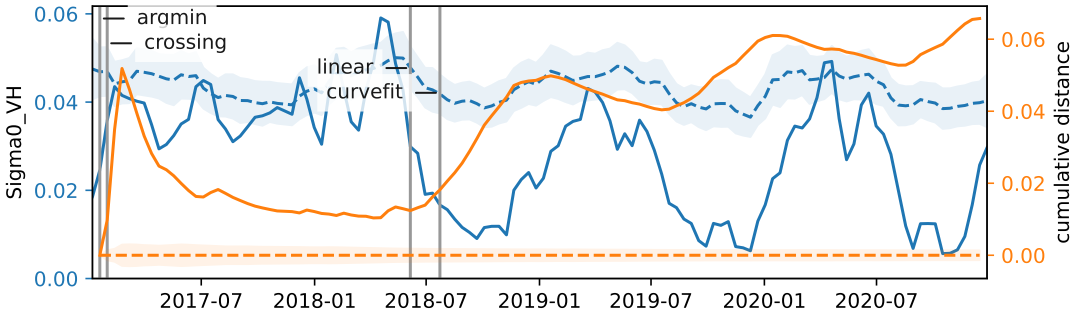
**Figure 5.20:** This figure illustrates the curve fit method for the cosine distance (fig. 5.20a) and the euclidean distance (fig. 5.20b). For the cosine distance, the change point placement requires a threshold on the S-curve, whereas the piecewise fit for the euclidean distance results in a sharp corner that gives a straightforward change point.

Looking at the relative performance of the different distance metrics and change detection methods, we can see some clear differences in the uniformity of the change maps, which could be an indicator for the robustness of the method. Overall, the *curve fit* and *linear* methods deliver the most visually appealing results with respect to the uniformity of the change map. This also means that the least amount of post-processing would be required for these methods if we wanted to vectorize these maps and extract change polygons.

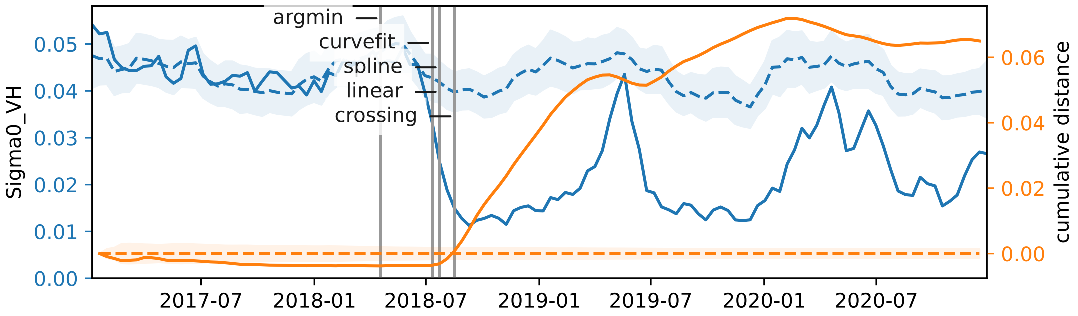
The *splines* method appears very conservative in the change point detection, resulting in a very patchy map. The *crossing* and *argmin* methods are also very patchy, but mostly as a result of false positives in the forest areas, and mis-predicting the date of change as very early in the time series when the true change only happens later. This is an indication that these methods are not very robust, as is expected: because of natural fluctuations, the minimum distance or first zero crossing may occur before the change event, and in particular may also be present in areas of stable forest.



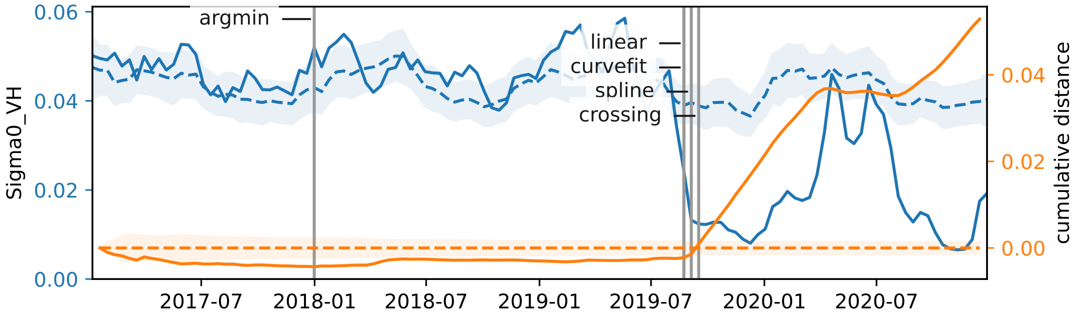
**(a)** Forest



**(b)** Non-forest

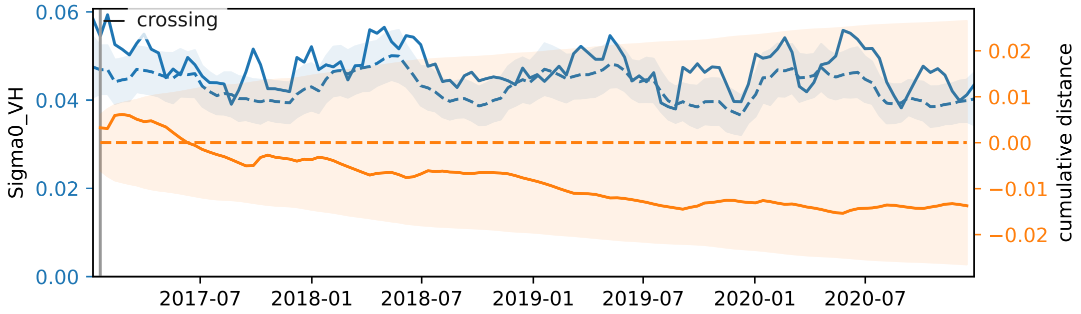


**(c)** Deforestation (pixel 1)

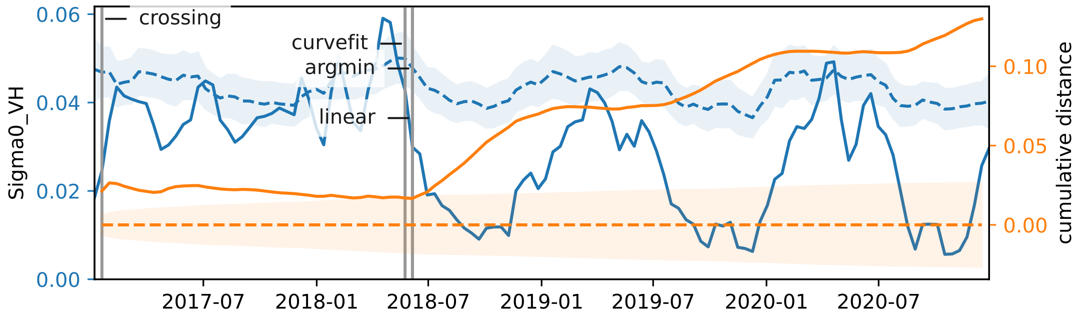


**(d)** Deforestation (pixel 2)

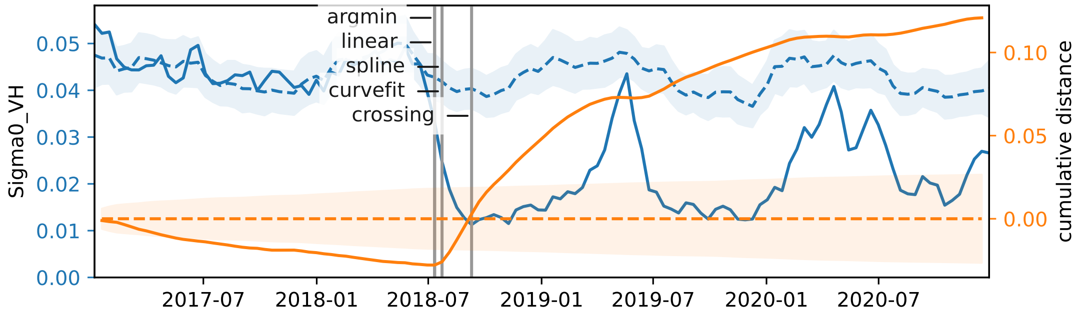
**Figure 5.21:** This figure shows the time series for four different pixels along with cumulative cosine distance, and the change points as detected by the five methods introduced in this section.



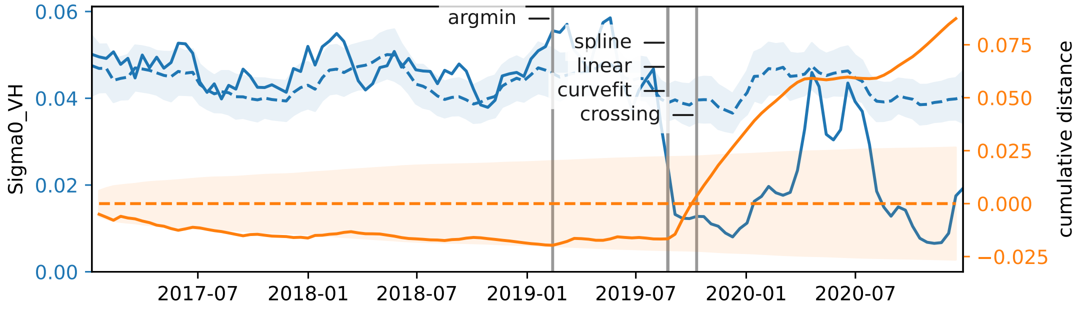
**(a)** Forest



**(b)** Non-forest

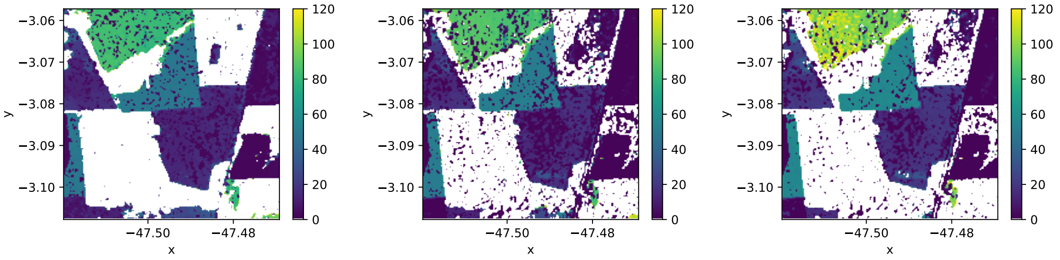


**(c)** Deforestation (pixel 1)



**(d)** Deforestation (pixel 2)

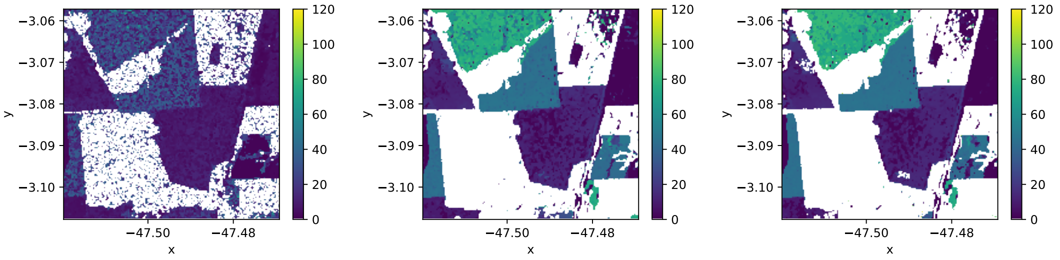
**Figure 5.22:** This figure shows the time series for four different pixels along with cumulative euclidean distance, and the change points as detected by the five methods introduced in this section.



cosine, crossing

euclidean, crossing

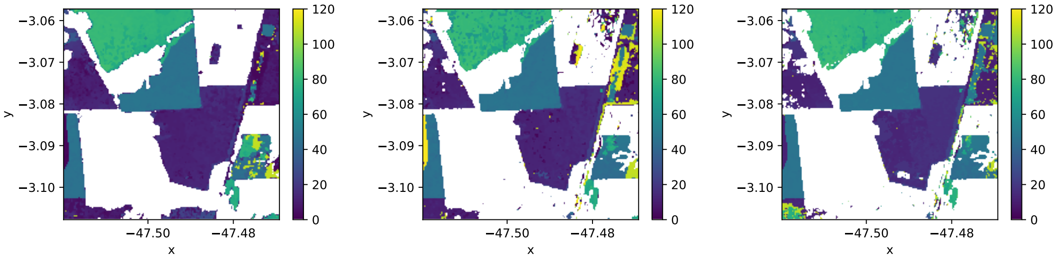
dtw, crossing



cosine, argmin

euclidean, argmin

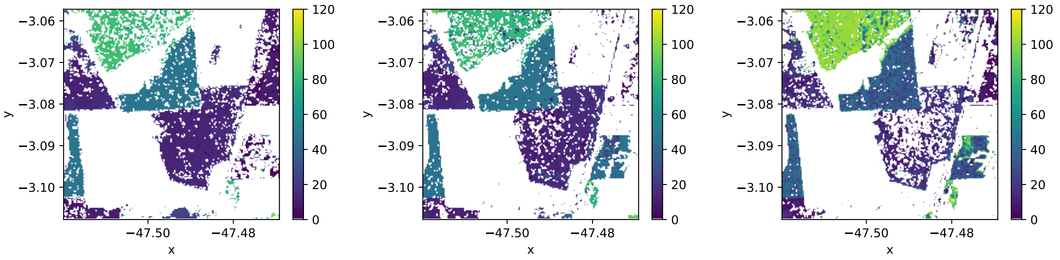
dtw, argmin



cosine, linear

euclidean, linear

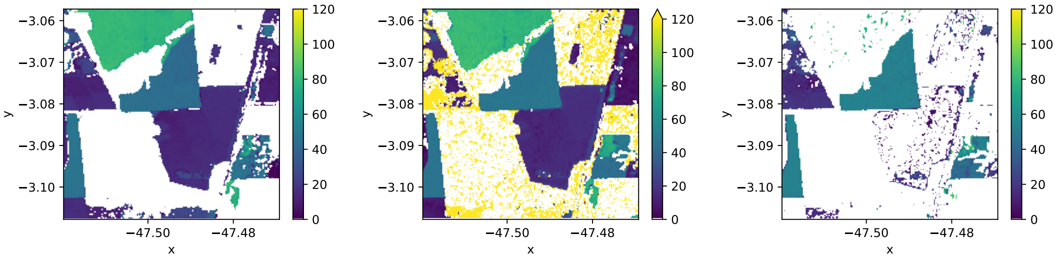
dtw, linear



cosine, spline

euclidean, spline

dtw, spline

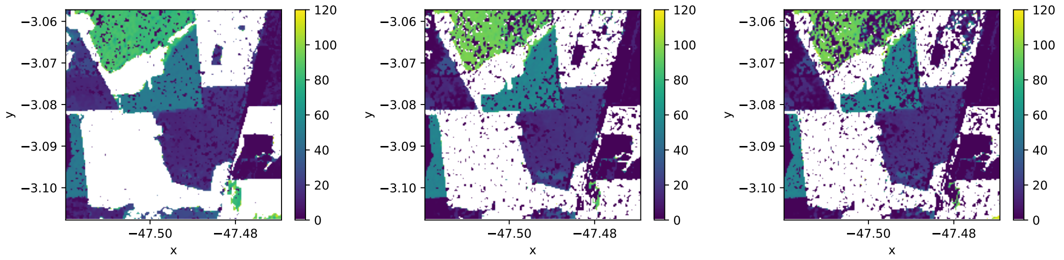


cosine, curve fit

euclidean, curve fit

dtw, curve fit

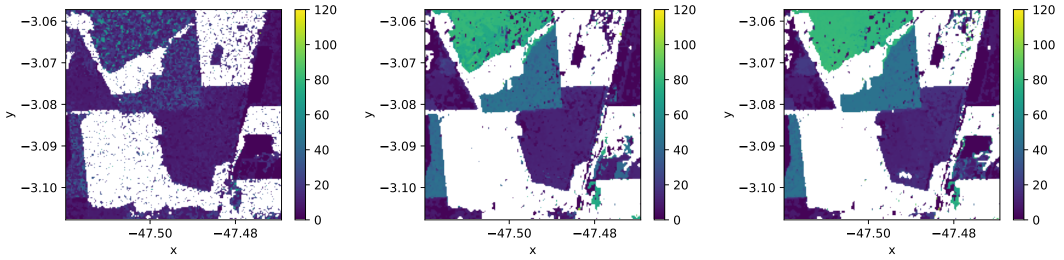
**Figure 5.23:** This figure shows the change maps based on the VH backscatter for three distance measures (euclidean, cosine, and DTW), and five different change detection methods (crossing, argmin, linear fit, spline fit, curve fit) as introduced in this section.



cosine, crossing

euclidean, crossing

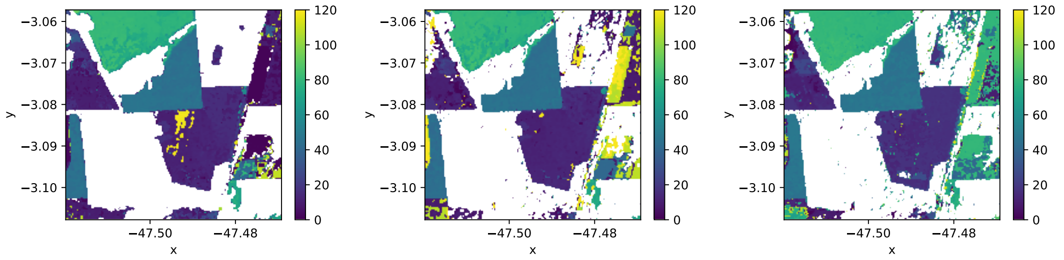
dtw, crossing



cosine, argmin

euclidean, argmin

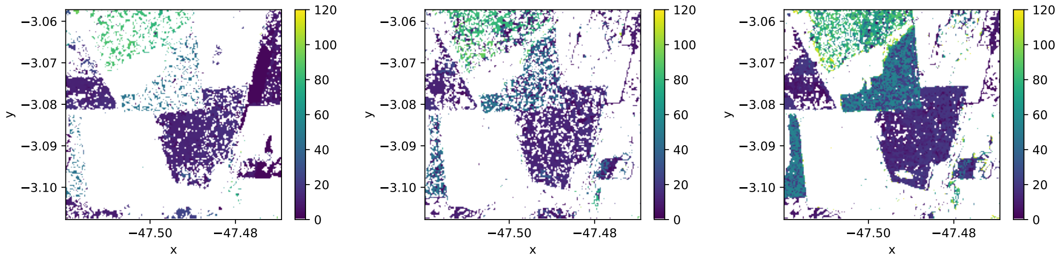
dtw, argmin



cosine, linear

euclidean, linear

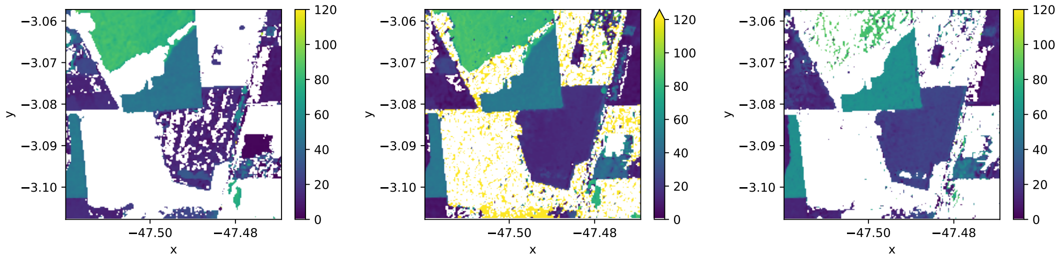
dtw, linear



cosine, spline

euclidean, spline

dtw, spline



cosine, curve fit

euclidean, curve fit

dtw, curve fit

**Figure 5.24:** This figure shows the same change maps as in fig. 5.23, except for the co-polarized VV backscatter.

The curve fit method appears to work very well for the cosine and euclidean distance, but applying the same method to the DTW distance does not yield similarly good results. Furthermore, this method works better for the cosine distance when the VH variable rather than the VV variable is used, whereas the opposite is true for the DTW distance.

In general, most methods have very low false positive rates in areas of intact forest, but high false positive rates in areas of non-forest. This indicates that a better technique is required for distinguishing existing non-forest areas from deforestation.

For each of the change maps in figs. 5.23 and 5.24 accuracy scores can be computed by treating the change detection as a classification problem. As the two reference data sets only provide yearly deforestation data, we consider only a binary classification into change/no change for the spatial accuracy (UA and PA). In addition, we assess the temporal accuracy as the mean lag of the detected date of change with respect to the change of date found through visual interpretation of Landsat imagery. These scores are listed in tables 5.1 to 5.3.

In general, most true deforestation events are reliably captured by many of the methods, resulting in very high PA scores. However, because many non-forest areas are mistakenly classified as deforestation, the UA values are expectedly low. Furthermore, in accordance with the observation that the splines method is more conservative in declaring a change point, it shows better UA values for most cases, but very low PA values.

Figure 5.25 illustrates the extracted change detection lag for the curve fit method with cosine distance. Red areas indicate a positive lag (the change is detected later than the true deforestation date), and blue areas indicate a negative lag (the change is detected earlier than the true deforestation date). Positive lags are expected and can represent either a low sensitivity of the respective change detection algorithm or a true delay in the backscatter response to the deforestation event. While the former reason can be addressed by improving the change detection algorithm, the latter is an inherent problem with using Sentinel-1 data for the purpose of near-real-time deforestation detection.

Negative lags, on the other hand, do not in fact represent the potential for early change detection as the change cannot physically be detected before it has occurred. Rather, it indicates either a mis-calibration of the detection algorithm (e.g.,

**Table 5.1:** Users's Accuracy for the change maps in figs. 5.23 and 5.24

| | variable | mode<br>metric | argmin | crossing | curvefit | linear | spline |
|---|---|---|---|---|---|---|---|
| Hansen | VH | cosine | 60.5% | 63.1% | 66.5% | 64.3% | 70.6% |
| | | dtw | 59.5% | 58.8% | 57.1% | 58.6% | 68.0% |
| | | euclidean | 62.3% | 61.2% | 60.6% | 61.3% | 73.5% |
| | VV | cosine | 56.7% | 62.8% | 65.5% | 64.6% | 61.0% |
| | | dtw | 61.1% | 55.9% | 59.7% | 61.3% | 76.3% |
| | | euclidean | 61.6% | 58.4% | 59.9% | 62.1% | 68.6% |
| JRC | VH | cosine | 40.5% | 44.8% | 48.7% | 46.5% | 52.1% |
| | | dtw | 46.0% | 43.9% | 41.4% | 46.6% | 52.3% |
| | | euclidean | 48.2% | 42.1% | 39.8% | 48.2% | 52.0% |
| | VV | cosine | 39.5% | 45.4% | 47.6% | 46.8% | 47.9% |
| | | dtw | 48.3% | 41.2% | 42.4% | 47.0% | 54.2% |
| | | euclidean | 46.0% | 41.7% | 39.4% | 45.0% | 53.5% |
| Visual interpretation | VH | cosine | 48.8% | 58.0% | 63.3% | 60.5% | 65.9% |
| | | dtw | 58.8% | 52.5% | 54.0% | 61.9% | 65.4% |
| | | euclidean | 61.6% | 56.9% | 55.4% | 58.9% | 68.3% |
| | VV | cosine | 53.6% | 59.3% | 61.5% | 60.1% | 61.2% |
| | | dtw | 59.0% | 53.2% | 56.4% | 58.3% | 68.1% |
| | | euclidean | 61.5% | 52.9% | 55.1% | 57.2% | 67.0% |

a threshold set too low in the curve fitting), or a false positive – an earlier event was mistakenly classified as the deforestation event, when the area was still in fact forest, such that the later true deforestation event is missed.

**Table 5.2:** Producer's Accuracy for the change maps in figs. 5.23 and 5.24

| | | mode | argmin | crossing | curvefit | linear | spline |
|---|---|---|---|---|---|---|---|
| | variable | metric | | | | | |
| Hansen | VH | cosine | 96.9% | 94.6% | 90.9% | 94.5% | 77.9% |
| | | dtw | 91.5% | 87.5% | 42.4% | 92.2% | 83.2% |
| | | euclidean | 91.4% | 88.5% | 92.7% | 90.5% | 78.8% |
| | VV | cosine | 96.9% | 94.1% | 81.3% | 93.7% | 49.4% |
| | | dtw | 90.5% | 90.4% | 75.1% | 89.9% | 70.5% |
| | | euclidean | 92.5% | 91.2% | 93.9% | 91.3% | 68.9% |
| JRC | VH | cosine | 98.8% | 98.0% | 95.5% | 97.8% | 84.4% |
| | | dtw | 94.3% | 92.3% | 36.1% | 93.8% | 72.4% |
| | | euclidean | 94.9% | 95.5% | 97.3% | 94.3% | 62.8% |
| | VV | cosine | 98.7% | 97.3% | 85.9% | 97.3% | 63.3% |
| | | dtw | 84.8% | 94.8% | 75.1% | 93.6% | 78.5% |
| | | euclidean | 95.9% | 95.1% | 96.8% | 95.5% | 66.2% |
| Visual interpretation | VH | cosine | 100.0% | 99.7% | 98.7% | 99.8% | 81.4% |
| | | dtw | 99.1% | 99.3% | 43.3% | 98.0% | 79.9% |
| | | euclidean | 99.4% | 98.2% | 99.8% | 99.5% | 80.2% |
| | VV | cosine | 99.8% | 99.5% | 89.0% | 99.7% | 62.5% |
| | | dtw | 98.7% | 98.8% | 80.3% | 99.0% | 94.1% |
| | | euclidean | 99.3% | 98.8% | 99.5% | 99.5% | 79.7% |

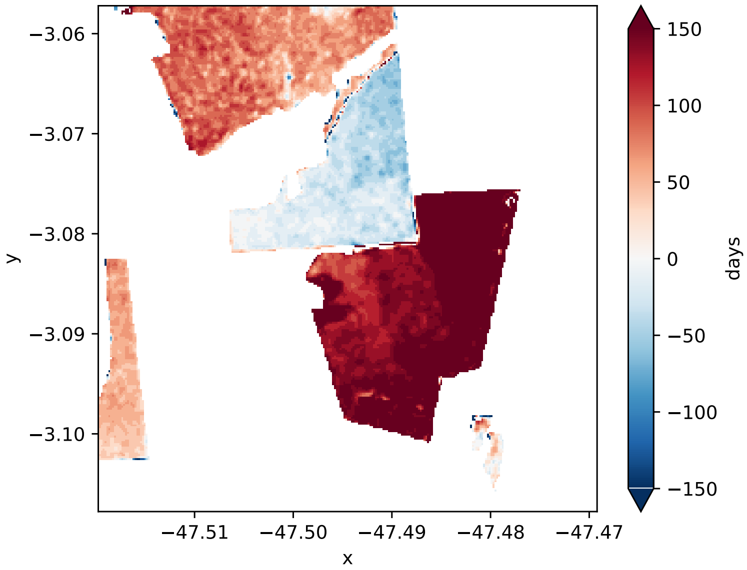


**Figure 5.25:** This figure shows the change detection lag per pixel in days, for the example of a curve fit change detection with a cosine distance. The lag is calculated with respect to the visually interpreted change dates.

**Table 5.3:** Mean lag for the change maps in figs. 5.23 and 5.24

| | | mode | argmin | crossing | curvefit | linear | spline |
|---|---|---|---|---|---|---|---|
| | variable | metric | | | | | |
| Visual interpretation | VH | cosine | -276.5 | 7.5 | 80.1 | 75.5 | 90.7 |
| | | dtw | 35.9 | 42.2 | 82.2 | 97.7 | 13.6 |
| | | euclidean | 20.9 | 110.0 | 99.4 | 76.0 | 50.7 |
| | VV | cosine | -202.6 | 95.5 | 90.3 | 102.0 | 91.6 |
| | | dtw | 36.2 | 48.2 | 183.8 | 116.8 | 1.0 |
| | | euclidean | 27.1 | 52.9 | 113.3 | 84.0 | -76.9 |

Figure 5.26 shows the mean time series for some of the polygons delineated in fig. 5.5. In addition, the mean change point is indicated for each polygon (using the curve fit method for the cosine distance), along with the standard deviation (marked as vertical gray band around the change point). The opacity of the marked change point and its uncertainty is scaled with the fraction of pixels in the area for which a change point was detected.

The first row represents stable forest. While changes are detected for some pixels in this area, e.g. due to imperfect polygon boundaries, the algorithm only produces very few false positives over forested areas, as visually indicated by the low opacity of the marked mean change point. On the other hand, the third row and the second to last row are examples of areas that are non-forest throughout the time period, but are consistently classified as deforestation events using the method shown. The third row in particular, and also the last row, are non-forest but look very similar to forest in their temporal radar signature, confusing the algorithm into mistakenly predicting deforestation.

In general, the change detection works very well in areas that do in fact undergo deforestation events (rows $2$, $4$, $5$, and $7$). The detected change point in these areas has low uncertainty and is generally close to the the point in the time series where we might manually place the point of change.

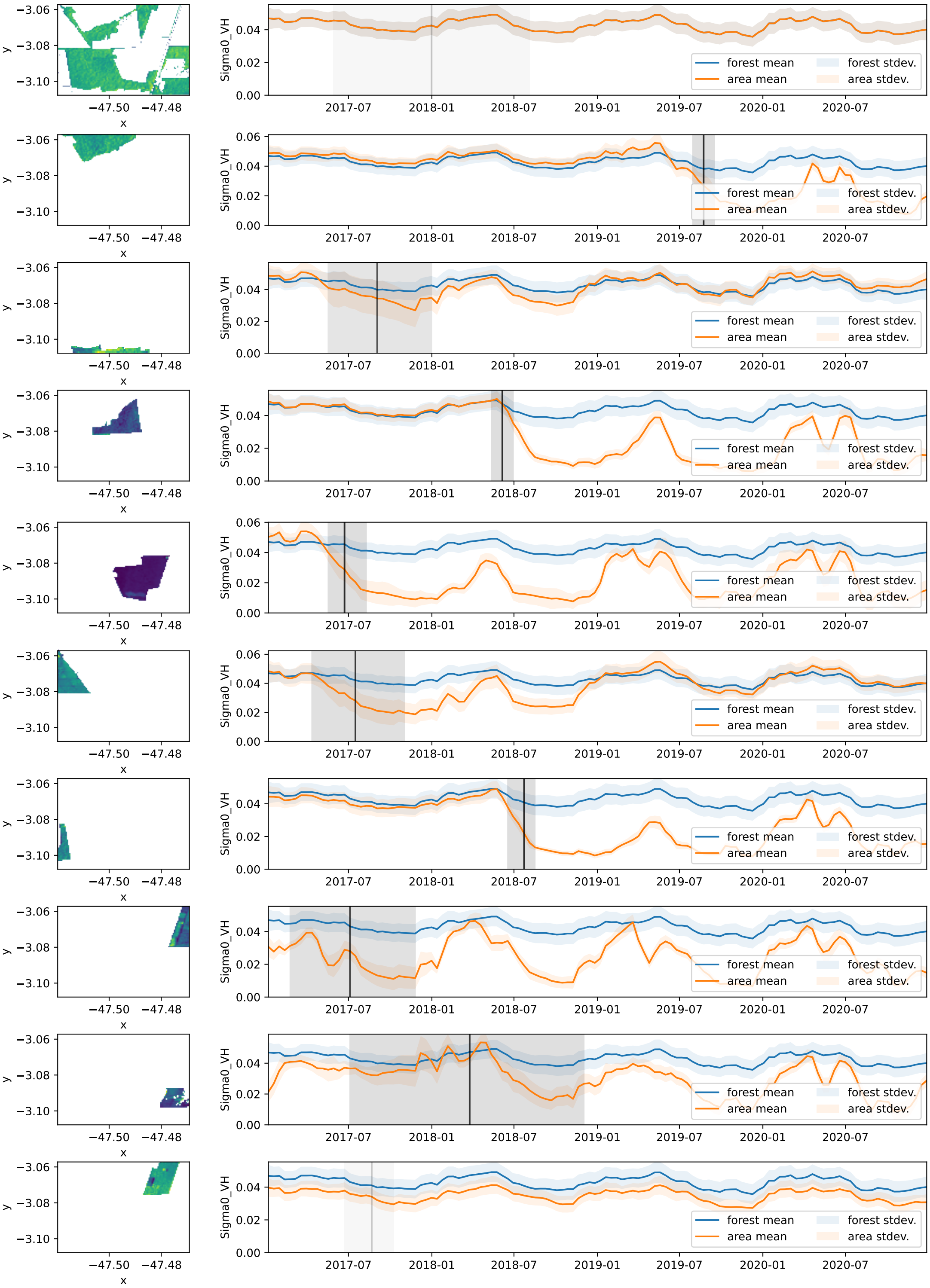
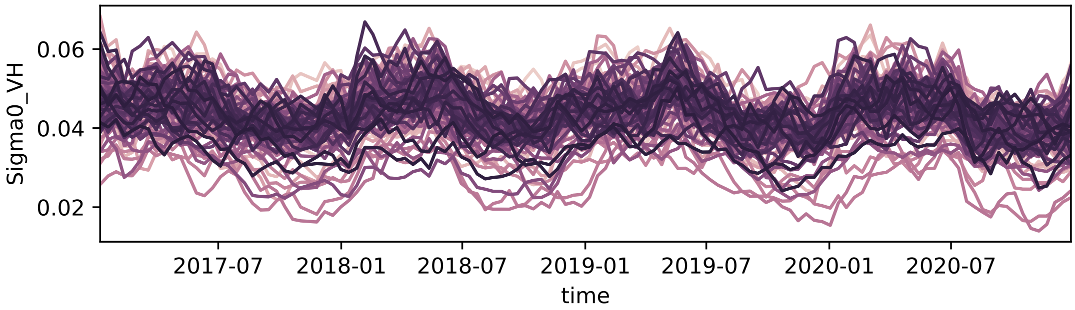


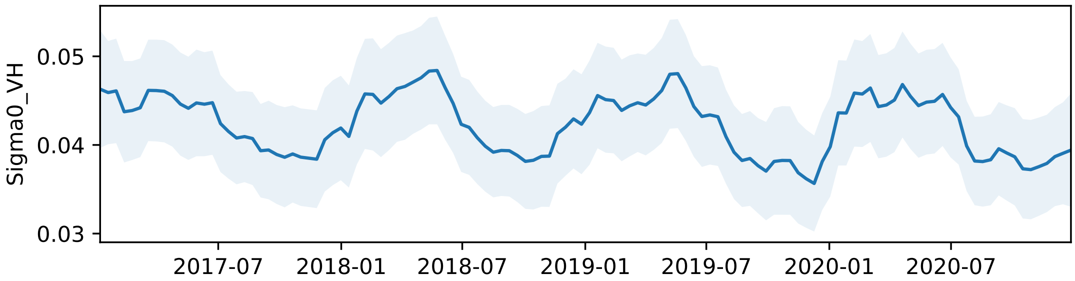
**Figure 5.26:** This figure shows the mean time series for some of the polygons shown in fig. 5.5 in orange. The blue line indicates the reference forest time series. Note that in the first row the forest time series coincides exactly with the time series of the area and is therefore hidden behind the orange line. The mean date of change for this area is shown as a vertical gray line (using the curve fit method with cosine distance). The gray band around this line indicates the mean plus/minus one standard deviation of all dates of change detected in the polygon on the left. The opacity of the change point line indicates the fraction of detected change points – e.g., for the first and last row, change points were only detected for a small number of pixels.

## 5.5 Forest Time Series Ensemble

For all results presented so far in this section, the forest distance of each pixel was calculated with respect to the *mean* forest time series. But what if the mean time series does not actually reflect the variety of temporal signatures that forest pixels can exhibit? To address this question, we will now be looking at many different forest time series to investigate how the inherent backscatter variance over forest impacts the change detection results. Figure 5.27 shows an ensemble of reference forest time series. These are obtained by applying the reference forest cover mask to the data set and selecting a random sample of pixels that are labeled as stable forest. The individual time series are shown in fig. 5.27a, while fig. 5.27b shows the mean and standard deviation.



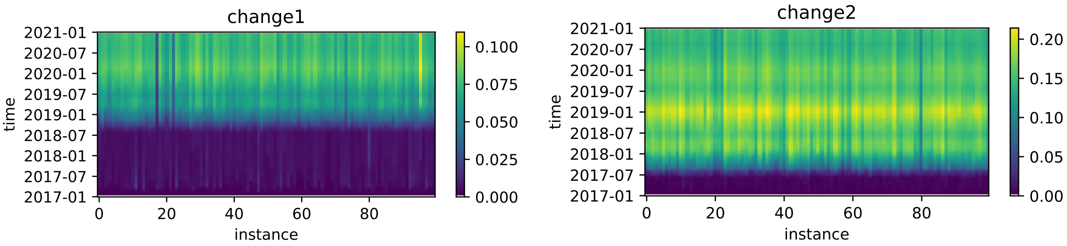
**(a)** Ensemble



**(b)** Ensemble mean

**Figure 5.27:** Forest ensemble

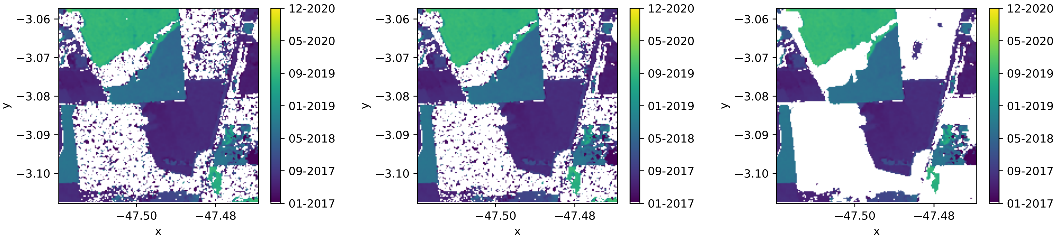
While we have previously computed the distance between each pixel and the mean forest time series (fig. 5.27b), we can also compute the distance with respect to each individual ensemble member (fig. 5.27) independently. For two sample pixels that are undergoing a change, this results in the change signatures displayed in fig. 5.28. These plots show the distance for every forest ensemble member over time. The pixel resembles some of these forest pixels more closely than others, and the color variability in the horizontal direction indicates this natural variability within forest. The time of change can be drawn through visual interpretation at the horizontal line of maximum contrast.



**(a)** Pixel 1 **(b)** Pixel 2

**Figure 5.28:** This figure demonstrates the change signature for two sample pixels given a reference forest ensemble made up of 100 individual time series. The color indicates the cumulative distance between the pixel and each ensemble member up to the corresponding time step. Visually, the time of change can be discerned as the horizontal line of maximum contrast.

To determine whether computing the distance separately for each ensemble member has a benefit for deforestation detection, we can perform the change detection for the cumulative distance computed for ensembles of different sizes. The effect of computing the distance separately for each ensemble member on the resulting change map is shown in fig. 5.29. Figures 5.29a and 5.29b show the behavior for ensemble sizes of 20 and 100, respectively, while fig. 5.29c shows the result when computing only the distance to the ensemble mean (corresponding to fig. 5.27b).



**(a)** $N = 20$ **(b)** $N = 100$ **(c)** Ensemble mean

**Figure 5.29:** Effect of ensemble size on change detection result

As we can see, the ensemble approach does not in general improve the change map, and in fact results in a larger number of false positives in forested areas. This is likely because the within-forest variability gets picked up as a significant deviation of some of the forest pixels from some of the other forest ensemble members.

In addition, the ensemble approach is more computationally expensive, as it requires a total of $TN_{\text{ensemble}}N_{\text{pixels}}$ distance calculations.

## 5.6 Effect of Mislabeled Data

Given that the idea behind this method is to be relatively independent of high quality training data, in this section the behavior of the method in the presence of mislabeled data is investigated. To simulate corrupted training data, a forest reference ensemble of size $N$ is generated for a range of fractions $c$ of incorrectly classified forest pixels such that $(1 - c)N$ ensemble members are true forest pixels, and $cN$ ensemble members are in fact non-forest. For each such corrupted reference ensemble, a change map was generated using the curve fit method with the cosine distance for the cross-polarized VH backscatter and the corresponding accuracy values were computed.

Figure 5.30 shows these maps of detected change for fractions between 0% and 28%. Up to about 10% incorrectly classified pixels, there is no major visible degradation of the change map. At higher percentages of corruption, changes that occur later in the time series are the first to be missed. However, a large amount of true changes are still detected up to 22% corruption. Furthermore, the corruption of the training data only increases the error of omission (the number of missed deforestation events), but not the commission error (the number of false positives). These findings are corroborated by a graph of the accuracy scores (UA, PA, and OA) as a function of the corruption fraction in fig. 5.31. The accuracy scores are computed with respect to the visually interpreted change reference (fig. 5.5). The PA is decreased significantly above 10% corruption, but the UA does not decrease – in fact, it increases slightly as the overall number of detected changes is reduced, including previous false positives. As a consequence, the OA does not deteriorate significantly even at 28% corruption. However, the PA is arguably the most important metric as the correct identification of all deforestation events is the primary goal. As stated previously, the false positives mostly occur in areas of non-forest and could be removed by other means, e.g. by comparing the results with optical data where non-forest areas may have been apparent even at the start of the time series.

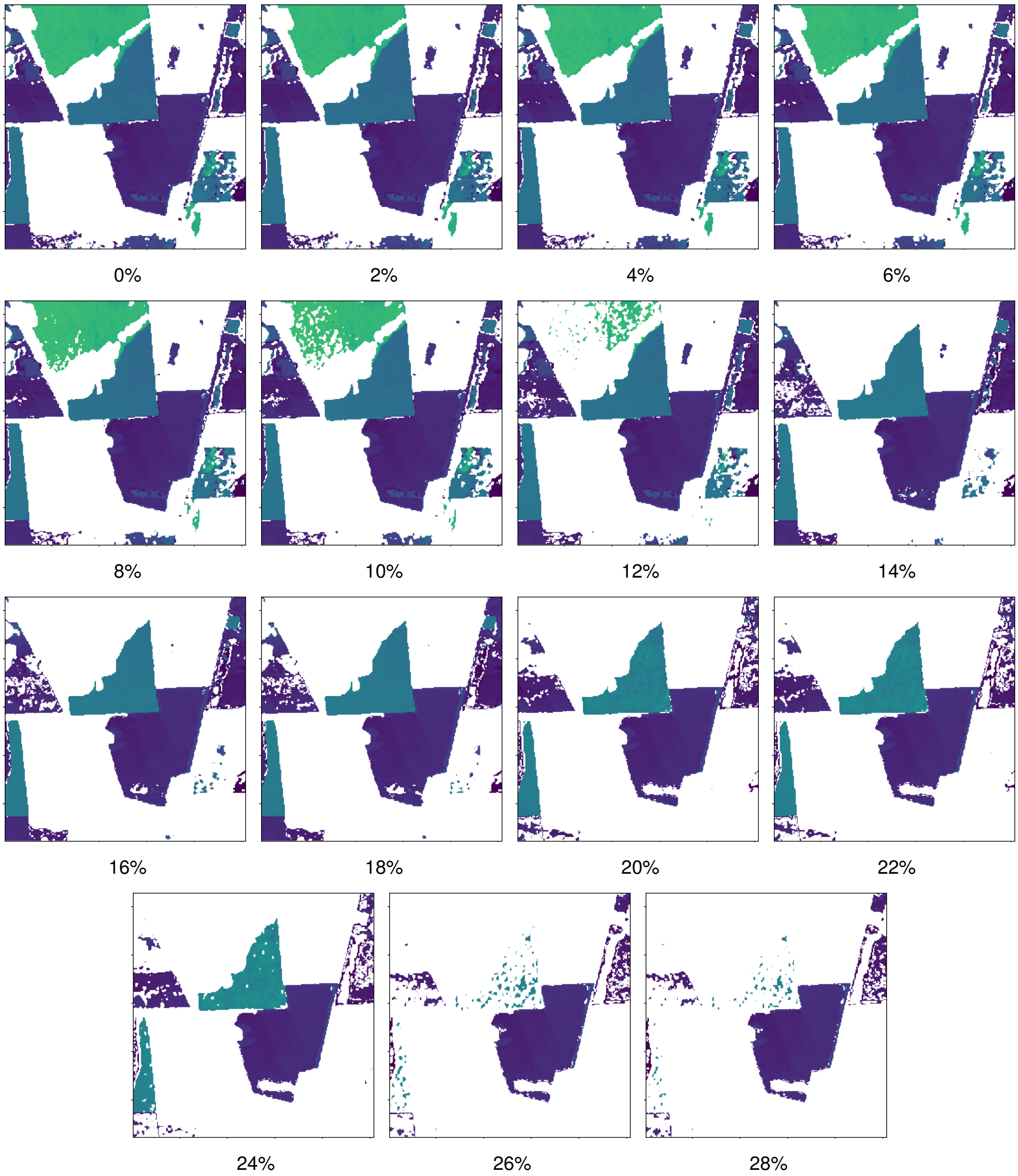


**Figure 5.30:** This figure shows change maps obtained from the curve fit method with a cosine distance and the VH backscatter (corresponding to the bottom left subfigure in fig. 5.23), for various levels of training data corruption. The percentages indicate the fraction of pixels in the forest reference mask that are in reality non-forest pixels.

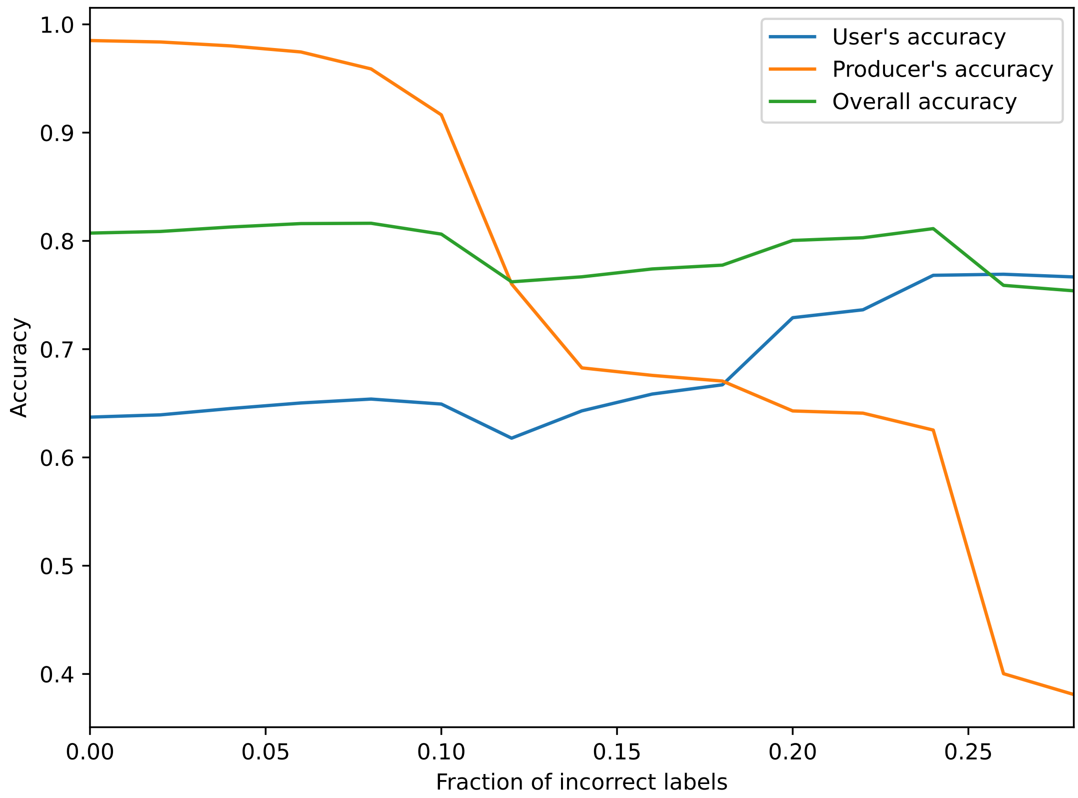


**Figure 5.31:** This graph shows the effect of training data corruption on the change detection accuracy. The UA, PA, and OA are shown as a function of the fraction of incorrect labels, using the curve fit method with a cosine distance and the VH backscatter.

## 5.7 Conclusions

In section 4.7, I listed some of the biggest challenges around deforestation detection using SAR data, in particular (1) the lack of reliable training data and (2) the fact that changes in backscatter need not represent changes in land cover/use. I have shown in this chapter that some of these challenges may be addressed with a semi-supervised approach where changes are detected with respect to the prototype time series of a reference class. This prototype time series captures the natural variability of the forest backscatter signature and removes the potential for mistakenly detecting changes due to other effects. In addition, this method works with noisy training data as it is robust to outliers present in the reference class ensemble, with no significant degradation of the results up to about 10% training data corruption. By detecting changes as deviations from a single reference class rather than by classifying between two possible classes, we remove the need to parametrize the non-forest class. This is advantageous for generalizing the method as the non-forest class is inherently heterogeneous and may include many different land cover types.

In conclusion, Sentinel-1 SAR data are certainly capable of detecting deforestation, even though a longer time series may be required to reliably tell apart forest and non-forest – potentially limiting the temporal change detection accuracy. There is no doubt that more innovations are possible in the field of multi-sensor fusion – combining the advantages of SAR, lidar, and optical data – and the use of spatial features for separating forest and non-forest with even shorter time series.

# Chapter 6

# An EO Python framework

Currently, there is a disconnect between the remote sensing and scientific computing communities. This is partly due to the discontinuity of the tools preferred in either community. The scientific computing community requires an API that integrates well with existing ecosystems for scientific data processing (I focus on Python here, but the same is true for other languages). In remote sensing, on the other hand, a large part of the processing is still done with graphical tools, e.g. SNAP or GIS software such as QGIS, ArcGIS, and ENVI. Most of these also provide a command line interface (CLI), as does the well-established tool GDAL.

However, when it comes to providing an API for programming languages, these tools lack in features or usability. GDAL for example does offer a Python API, but it is clear that this API has been developed as a wrapper for the CLI rather than envisioned around concepts more native to the programming language. This shows e.g. in the fact that most operations that can be performed in the Python GDAL API still read one file and output another, rather than operating on in-memory objects. The popular Python library `rasterio` (https://github.com/mapbox/rasterio) is a more Python-friendly wrapper around GDAL, but is still cumbersome when working with more than a single image.

There is essentially no library out there in any programming language that natively implements e.g. pre-processing of SAR data without having to go through the CLI of some GUI based software. Also, operations such as reprojection, mosaicking and tiling are not easily available outside of dedicated command line tools. All of this means it is not easy for a scientist or developer to implement a new algorithm without worrying about IO, file formats, and preprocessing because there is no library that implements all of those steps in a sufficiently modular way to allow for the addition of new algorithms as modules (other than as plugins for GUI software). Python developers need libraries that neatly fit into the SciPy ecosystem (https://www.scipy.org/) which provides the basis for all scientific computing in Python. Having established the need for a *grand unified EO library*, this chapter aims to present some ideas as to how such a library may be structured.

Another issue with Earth observation data processing is the amount of data required. For applications dealing with global datasets it is often unfeasible to download all data locally for processing. Even if storage, memory, and compute resources are not an issue, bandwidth limitations mean that obtaining such datasets could be a matter of months or years. Projects like *Google Earth Engine* (Gorelick et al. 2017) or *AWS Earth* aim to overcome this limitation by providing datasets that exist physically right next to their cloud computing services such that no data will need to be transferred via low bandwidth connections until the export of the final product. In most cases, the generated data product is much smaller than the sum of all the raw input data. It took Google Earth Engine more than three years to acquire the multi-petabyte dataset which is now available, partly because the Landsat archive had to be transferred from the long-term tape storage (Gorelick et al. 2017). While the focus here is not on cloud services like this, a well-designed library with scalable tools would also be suitable to run on such a service.

Perhaps the largest barrier is the lack of a common standard file format. While GeoTiff is the de-facto standard for most applications exchanging geospatial raster data, it has a number of crucial drawbacks. Firstly, it isn't capable of storing data in more than two dimensions and as such is unsuitable for datacube structures. Secondly, while GeoTiff files do contain geocoding information (that is what makes it a GeoTiff rather than a Tiff), they are far from self-descriptive: they don't support storing band names, coordinate system information, etc. When software such as QGIS opens a GeoTiff, they will immediately create an additional XML file storing metadata such as projection and display parameters that are gathered from user input. Clearly, GeoTiff is not fully up to the task of geospatial data processing if it requires a separate metadata file to work properly.

Various libraries such as satpy (Raspaud et al. 2022), eo-learn (Aleksandrov et al. 2022), RSGISLib (Bunting et al. 2014), and Pyeo (Roberts et al. 2020) have been created in an attempt to solve some of these problems, often with a particular focus on machine learning. However, all of these libraries work with two-dimensional image data rather than operating on data cubes that inherently include a temporal data dimension.

Rather than inventing yet another file format or even a new Python data structure, the project builds on the `xarray` library (Hoyer et al. 2017) which interfaces neatly with the NetCDF file format and essentially extends `pandas` to $n$-dimensional data. `xarray` has been developed by the climate data community and is now officially recommended by `pandas` in favor of the deprecated `pandas.Panel` (https://pandas.pydata.org/pandas-docs/version/0.23.4/generated/pandas.Panel.html).

The library I have developed is called `nd` to reflect the central role of $n$-dimensional data in this framework, i.e. breaking out of the restrictions of two spatial dimensions imposed by GeoTiff and related software. The current version of this library is `v0.3.1` (as of 1 Nov, 2021). It is still constantly being developed and originally arose from the code needed for generating the results presented in this thesis. The code is hosted at https://github.com/jnhansen/nd, published on Zenodo (Hansen 2021), and also available from Python Package Index (PyPI): https://pypi.org/project/nd/. Unit tests are run on Travis CI for Python versions 3.6 to 3.8 with a code test coverage of 91%. The full documentation may be accessed at https://nd.readthedocs.org/.

## 6.1 The NetCDF File Format

NetCDF (specifically NetCDF-4) is a highly efficient file format that was built on top of HDF5. It is capable of random access which ties in with indexing and slicing in `numpy`. Because slices of a large dataset can be accessed independently, it becomes feasible to handle larger-than-memory file sizes. NetCDF-4 also supports data compression using `zlib`. Random access capability for compressed data is maintained through data chunking. Furthermore, NetCDF is designed to be fully self-descriptive. Crucially, it has a concept of named dimensions and coordinates, can store units and arbitrary metadata.

While NetCDF has been the file format of choice in climate modeling for a long time, the Earth observation community is only recently adopting it. To be convinced of the slow but steady adoption of NetCDF it is revealing to look at the evolution of file formats used by ESA for the Copernicus missions. Sentinel-1 and Sentinel-2 still use the proprietary file format SAFE which is basically a collection of GeoTiff files put together in a folder with metadata. Sentinel-3 uses the SEN3 format, which works similarly to SAFE, except now it is a collection of NetCDF files put together in a folder with metadata. Finally, with Sentinel-5p, ESA has made the move to a single NetCDF file for each product.

In order to facilitate the work with existing file formats, the framework aims to provide an interface to read and write file formats to/from NetCDF via its Python cousin `xarray`. The compatibility layer between other file formats and NetCDF is mostly fed by GDAL via `rasterio`, which offers good support for a long list of geospatial file formats (https://www.gdal.org/formats_list.html).

## 6.2 Package Structure

Figure 6.1 shows an architecture for a framework to achieve this. In essence, there is a data compatibility layer which handles the ingestion of geospatial data formats into `xarray`, via `rasterio` and GDAL. As mentioned previously, the library does not aim at creating yet another data structure, but rather builds around the multivariate datacube as represented by an `xarray` Dataset.

The second important part of the library is the algorithms. The library provides an abstract base class `Algorithm`, which contains the basic scaffolding for any algorithm that may be added as a module. This includes basic parallelism and capabilities for distributed computing, e.g. using Python `dask` (https://dask.org/). Parallelism is even extended to account for edge effects by adding a buffer around individual chunks. It further ensures that all algorithms follow the same basic structure in terms of parameters and outputs to ensure consistency across all modules and to reduce the learning curve for new developers and users. This is loosely modeled after the way `scikit-learn` (http://scikit-learn.org/) implements its algorithms.

A number of basic algorithms need to be implemented before the library becomes attractive for researchers. However, after that threshold the modular structure built on existing tools makes contributing easy and will invite algorithm contributions from other scientists.

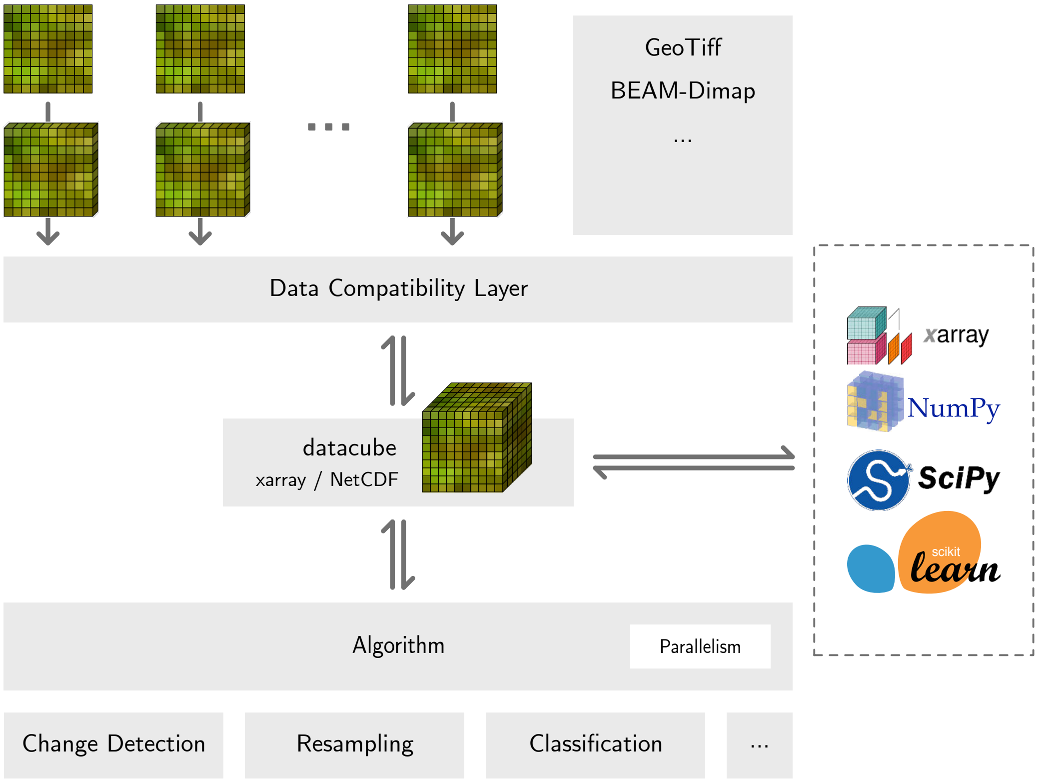


**Figure 6.1:** This chart illustrates what a Python library for EO data processing might look like.

The package is currently structured according to the following (main) submodules:

**`nd.change`** contains various change detection algorithms.

**`nd.classify`** provides an interface to apply `scikit-learn` classifiers to `xarray` Datasets.

**`nd.filters`** contains a selection of filters for $n$-dimensional datacubes, e.g. boxcar, Gaussian, arbitrary kernel convolutions, and non-local means.

**`nd.io`** collates all methods to read from and write to various file formats, largely wrapping `xarray` and `rasterio` methods.

**`nd.testing`** provides extra methods used by the unit tests.

**`nd.tiling`** allows for splitting of a dataset into tiles as well as merging tiles into a single dataset, including buffered tiles.

**`nd.utils`** is a collection of various utility functions used by other submodules.

**`nd.vector`** provides tools to work with vector data, in particular to rasterize such data to match a gridded reference dataset.

**`nd.visualize`** provides methods to create visualizations of xarray datasets, e.g. maps, RGB images, and video.

**`nd.warp`** contains functionality around reprojections between coordinate reference systems, resampling, etc.

The following sections are meant to provide a user guide and a more detailed overview of the currently implemented features.

## 6.3 Installation

Prior to installing the library you may also want to install the GDAL library, even though `rasterio` comes with a stripped down version of GDAL so for most use cases this should not be necessary.

Note that I do not currently officially support or test installation on Windows, but people have had mixed results. The easiest way to install `nd` is via `pip` from PyPI:

```
pip install numpy
pip install nd
```

You can also install the latest version from Github:

```
pip install git+https://github.com/jnhansen/nd
```

In addition, the following algorithms require the `libgsl-dev` C library:

- `nd.change.OmnibusTest`

If you want to use these algorithms you need to make sure you have the library installed before installing `nd`. You can find out whether it is installed by checking if the command `gsl-config` exists on your machine.

### Rebuilding the C extensions from Cython

In case you want to rebuild the C extensions from the `.pyx` files, you need to install the additional dependencies `cython` and `cythongsl`. With those installed, `pip install` will automatically regenerate the C files prior to compilation.

## 6.4 Examples

### 6.4.1 Data visualization

When you want to plot a single variable for a single time slice, `xarray`'s built-in plotting methods are more than sufficient. `nd` provides a few extra methods for visualizing multi-temporal and multivariate datasets.

```
import nd
ds = nd.open_dataset('data/C2.nc')
t0 = ds.isel(time=0)
rgb = nd.to_rgb([t0.C11, t0.C22, t0.C11 / t0.C22],
                'images/c2_rgb.png')
```

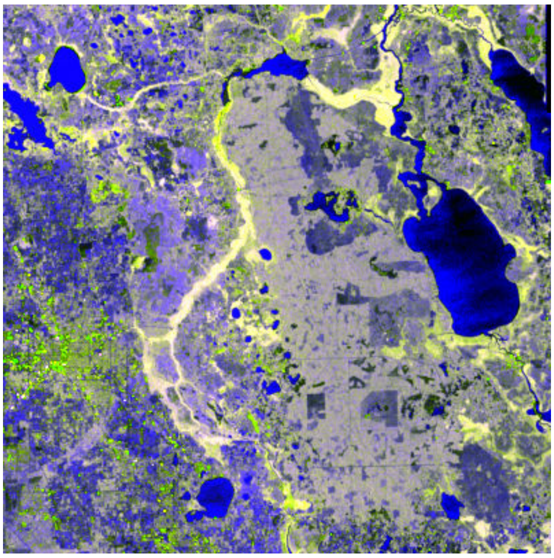


**Figure 6.2:** RGB representation

**Creating video from time series data** The example writes to a gif image so it can be embedded here, but many video formats are supported.

```
from nd.visualize import write_video
write_video(ds, 'images/c2.gif', fps=5, timestamp=False)
```

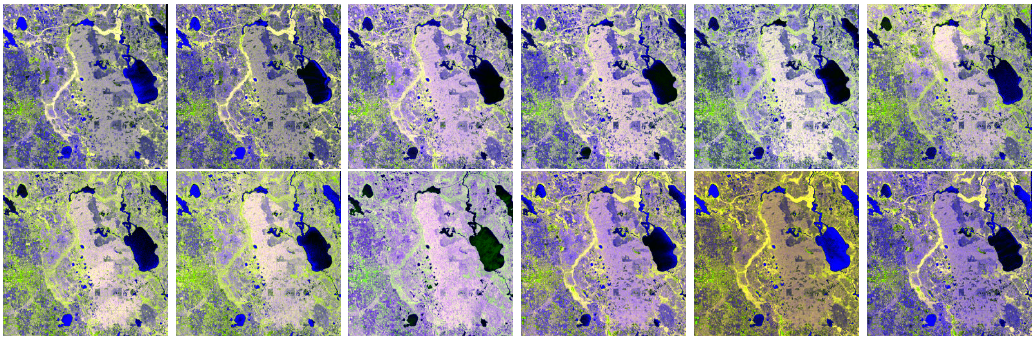


**Figure 6.3:** The individual frames of a video created using `nd.visualize.write_video()`.

**Creating a map** This example creates a simple map showing the outline of the dataset as a polygon.

```
from nd.visualize import plot_map
import matplotlib.pyplot as plt
plt.figure()
plot_map(ds, buffer=6, imscale=11)
```

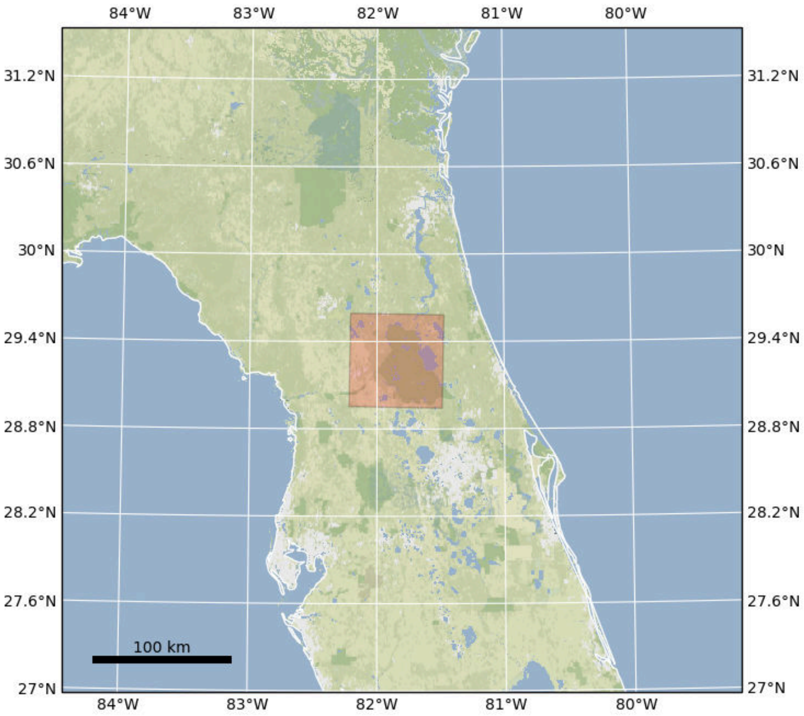


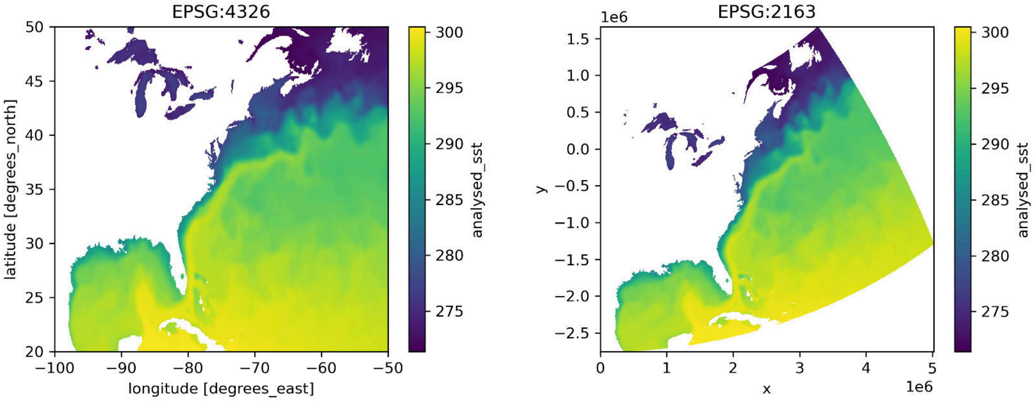
**Figure 6.4:** Representation of the dataset geometry on a map

### 6.4.2 Reprojection

Reprojecting a dataset into a new coordinate reference system can be easily achieved using `ds.nd.reproject()`. This function is a high-level wrapper around `rasterio`'s own reprojection methods and maintains the dataset structure including additional (non-spatial) dimensions. The following example demonstrates the reprojection on some GHRSST data (JPL OurOcean 2010).

```
ds_proj = ds.nd.reproject(
    src_crs='epsg:4326',
    dst_crs='epsg:2163')
```

Figure 6.5 shows the dataset before and after reprojection.



EPSG:4326 EPSG:2163

**Figure 6.5:** Reprojection of a subset of GHRSST data from EPSG:4326 (WGS84) to EPSG:2163 (US National Atlas Equal Area)

### 6.4.3 Applying a function across arbitrary dimensions

The following code snippet demonstrates how a function of arbitrary signature acting on any subset of dimensions can be easily mapped across a dataset using `ds.nd.apply()`. In this example we are reducing the time dimension by mapping the function `np.argmax` to find the hottest month for each pixel. While this example is quite trivial and could be solved without using `ds.nd.apply()`, the method is very flexible and can efficiently map functions of any signature over arbitrary dimensions.

```
hottest = ds.nd.apply(
    np.argmax,
    signature='(time)->()')
```

Figure 6.6 shows the result for this function call, displaying the hottest month per pixel for the GHRSST dataset.

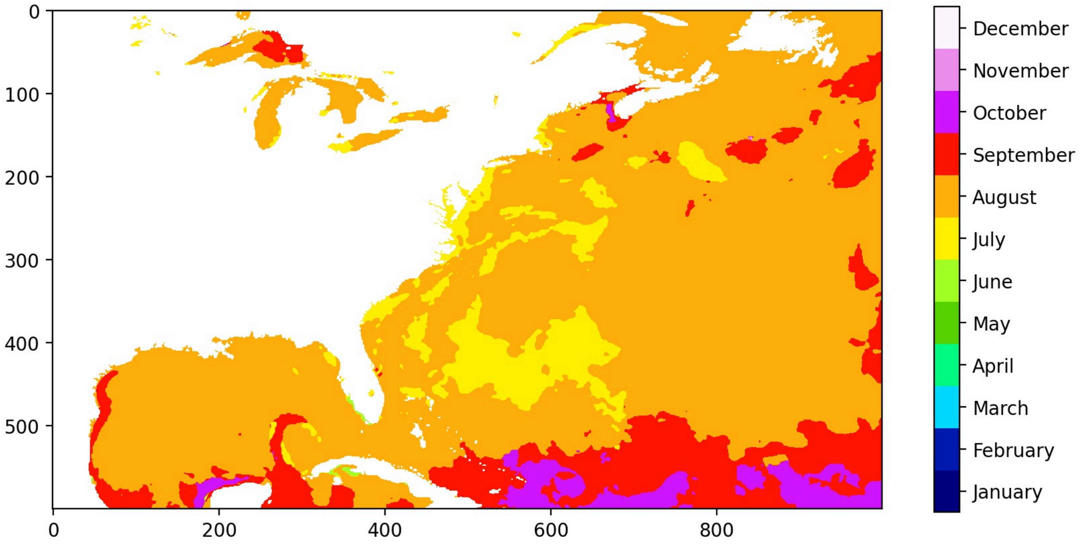


**Figure 6.6:** This figure shows the hottest month for each pixel in this subset of the GHRSST data. The result was obtained by mapping `np.argmax` over the dataset using `ds.nd.apply()`.

### 6.4.4 Integration with `scikit-learn`

`nd` provides a wrapper for `scikit-learn` classifiers that can be directly trained on and applied to `xarray` datasets.

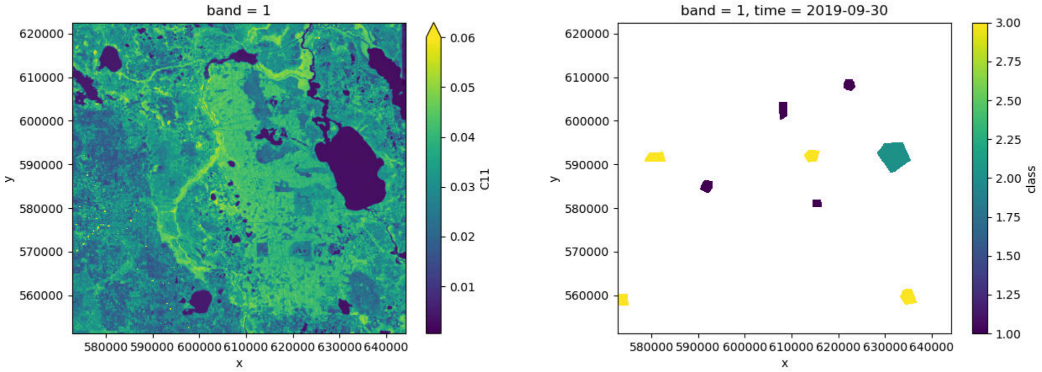
First, we need to get some training data. We will do a forest/non-forest classification using a Sentinel-1 time series over an area in Florida. We also have some polygon training data, which can be rasterized to match the dataset as follows:

```
labels = nd.vector.rasterize('labels.shp', ds)
```

If we investigate `labels`, we see that it has an associated legend to match the integer classes:

```
>>> labels
<xarray.DataArray 'class' (y: 400, x: 400)>
array([[0, 0, 0, ..., 0, 0, 0],
       [0, 0, 0, ..., 0, 0, 0],
       [0, 0, 0, ..., 0, 0, 0],
       ...,
       [0, 0, 0, ..., 0, 0, 0],
       [0, 0, 0, ..., 0, 0, 0],
       [0, 0, 0, ..., 0, 0, 0]])
Coordinates:
    band     int64 1
```

```
  * y        (y) float64 6.225e+05 6.223e+05 6.221e+05 ...
                          5.515e+05 5.513e+05
  * x        (x) float64 5.729e+05 5.73e+05 5.732e+05 ...
                          6.438e+05 6.44e+05
    time     datetime64[ns] 2019-09-30
Attributes:
    legend:   [(0, None), (1, 'forest'), (2, 'water'),
               (3, 'nonforest')]
```



**(a)** Data

**(b)** Classification labels

**Figure 6.7:** This figure shows the input data (Sentinel-1) (fig. 6.7a) along with the rasterized polygon labels (fig. 6.7b).

To apply a `scikit-learn` classifier to an `xarray` dataset we can simply wrap it in a `Classifier` provided by `nd.classify`:

```
from nd.classify import Classifier
from sklearn.ensemble import RandomForestClassifier
clf = Classifier(RandomForestClassifier(n_estimators=10))
pred = clf.fit(ds, labels).predict(ds)
pred.isel(time=0).plot()
```

If we plot the mean of the predicted class over time we can see that the predictions change because the forest cover changes over the course of the time period:

```
>>> pred.mean('time').plot()
```

**Clustering** Clustering can be done using the same `Classifier` object because clustering classes in `scikit-learn` provide the same interface as classifiers. Clustering is an unsupervised approach, so the `labels` argument will be omitted.

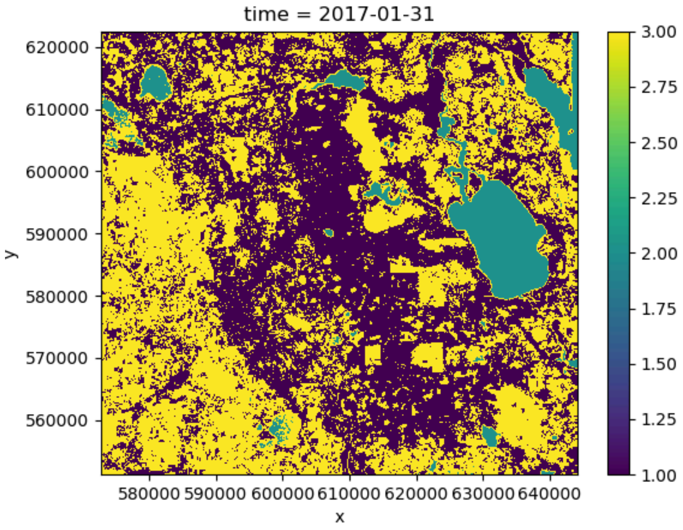


**Figure 6.8:** This figure shows the predicted class labels for one time step.

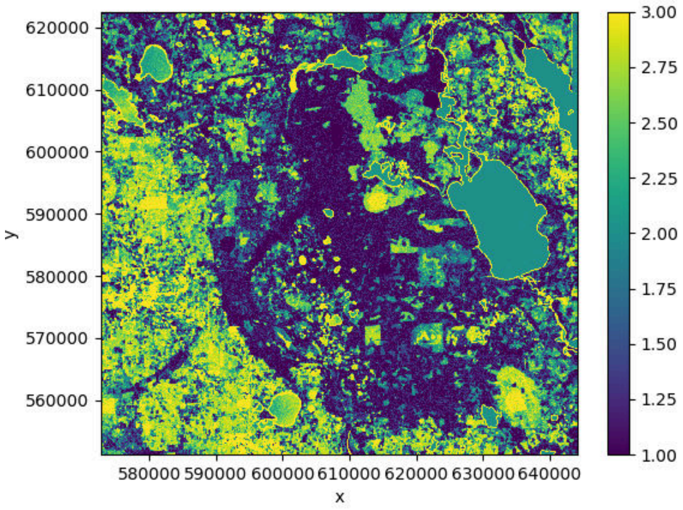


**Figure 6.9:** This figure shows the mean predicted class label over time.

In the following example, we are using `nd.classify.class_mean()` to replace every pixel with the mean of its cluster for visualization:

```
>>> from sklearn.cluster import MiniBatchKMeans
>>> from nd.classify import class_mean
>>> clf = Classifier(MiniBatchKMeans(n_clusters=3))
>>> pred = clf.fit_predict(ds.isel(time=0))
>>> means = class_mean(ds.isel(time=0), pred)
>>> means.C11.plot()
```

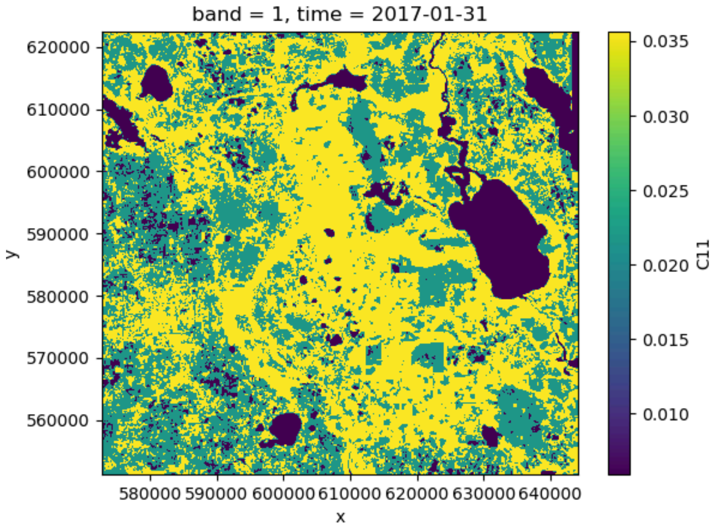


**Figure 6.10:** This figure demonstrates a sample clustering outcome with $k = 3$ clusters.

It is advisable to use a clustering algorithm that scales well, such as `MiniBatchKMeans`. Alternatively, one can fit the clusterer to a smaller subset of the data by applying a mask.

**Feature and data dimensions** Internally, the entire dataset needs to be converted to a two-dimensional array to work with most classification algorithms in `scikit-learn`. The first dimension (rows) corresponds to independent data points, whereas the second dimension (columns) corresponds to the features (attributes, variables) of that data point. By default, `nd` will flatten all dataset dimensions into the rows of the array, and convert the data variables into the columns of the array.

However, `nd.classify.Classifier` has an additional keyword argument `feature_dims` that controls which dimensions are considered to be features of the data. Typically, this could be a `band` dimension, which really isn't a dimension but a set of features (or variables) of the data. It could also be the `time` dimension, in which case all time steps are treated as additional information about a point, rather than separate points in the feature space.

Example:

```
>>> clf = Classifier(
...     RandomForestClassifier(n_estimators=10),
...     feature_dims=['time'])
>>> pred = clf.fit(ds, labels).predict(ds)
```

```
>>> pred.plot()
```

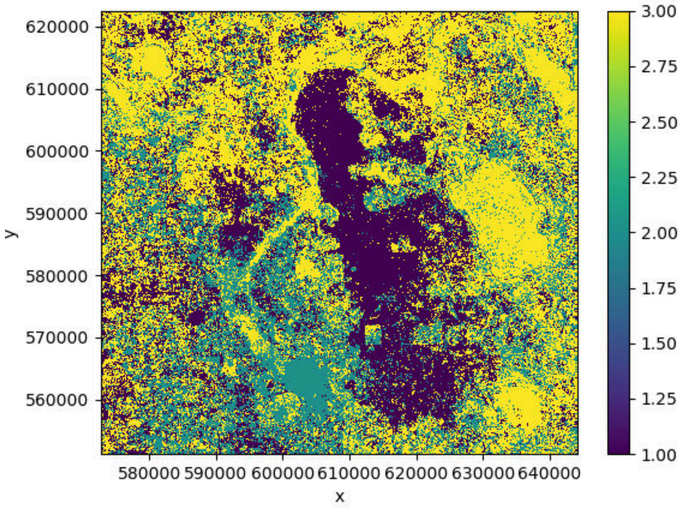


**Figure 6.11:** This figure shows the predicted class labels when the time dimension is used as features.

Our prediction output no longer has a `time` dimension because it was converted into a feature dimension and used for prediction. In this case the result is not great because the classes change over time and as a consequence we have noisy training data.

## 6.5 Conclusions

I have created the `nd` library while working on the various data analysis and algorithm development projects during my PhD.

The library integrates the Python libraries `scikit-learn`, `rasterio`, and `xarray`, among others, to seamlessly perform common classification and reprojection tasks on `xarray` datasets and therefore NetCDF files. It is built around the central concept that datasets may have an arbitrary number of dimensions, such as spatial and temporal dimensions, multispectral bands, etc. The algorithm base class ensures that all provided algorithms work on arbitrary dimensions and can be easily parallelized.

It now contains over $7000$ lines of code, mostly written in Python, but with some high performance code in Cython for compilation into C. It is primarily hosted on

GitHub (https://github.com/jnhansen/nd) and is published on the Python Package Index (PyPI) with the current version `v0.3.1`.

I am expecting to continue maintaining and developing this library as I continue writing code for remote sensing applications. Pull requests and issues on GitHub are always welcome.

# Chapter 7

# Discussion

In this thesis I have aimed to tackle the problem of timely deforestation detection, with a specific focus on avoiding reliance on high quality reference data, while remaining specific to deforestation as opposed to other changes.

I chose to focus on SAR data due to its independence of cloud cover and therefore high temporal observation density, making it ideal for near-real time monitoring. I discussed the theory and motivation behind SAR imaging as well as its application to deforestation monitoring. I reviewed existing change detection algorithms for both optical and SAR data as well as mechanisms to combine multiple data sources. A common problem among these algorithms is the general absence of spatial context in the models. Not considering the spatial neighborhood of a pixel means disregarding the large potential that lies in data cubes.

I briefly assessed interpolation errors caused by reprojection from the native image coordinates to a geospatial projection. It turned out that interpolation likely causes geographic registration errors of up to $7$ m in both latitude and longitude in the areas between GCPs. This is in the order of approx. 1 pixel for an SLC product. While likely not an issue for coregistration, this is certainly an effect to be aware of.

The theory predicts the limitations of C-band SAR for deforestation monitoring. With its $5.5$ cm wavelength, Sentinel-1 fails to penetrate dense forest canopy. At this wavelength, all interactions between the radar and objects on the ground happen in the top layer of the canopy, such that the backscatter contains little to no information about the forest floor or understory where most of the biomass is located. While this is certainly a problem for accurately and quantitatively estimating biomass, clearcut deforestation can well be detected using this wavelength range, especially as a complement to additional data sources.

In chapter 3 and in particular in section 3.3.1, I looked at the separability between forest and non-forest using different SAR parameters (backscatter and phase) obtained from Sentinel-1 SLC data. In general, the separability of these two classes was greatly dependent on the study site. The differences between study sites suggest that the accuracy of time series based classification is highest in areas of pronounced seasonality, as witnessed in, e.g., Finland or the UK, where accuracies of above 90% were achieved. Nevertheless, even in areas with low seasonality such as the tropics, accuracies of around 80% to 85% paired with the independence of SAR data from cloud cover make a strong case for its use. A larger selection of study sites may be required to determine whether these differences are due to the characteristics of the individual study sites or a feature of the broader climate zones, forest types, and types of forest disturbance found in each site. As a general result, forest was characterized by larger backscatter values in both VV and VH, as well as larger annual variation in VV and VH compared to non-forest. There was little information found in the phase difference and no information in the phase variability. The UK study site exhibited a curious shift in phase distributions between forest and non-forest for one particular orbit direction, which I was able to explain as an artifact of the underlying topography. These results appear to suggest that there is no reason to use SLC over GRD data as there is no gain in useful information. However, this conclusion needs to be qualified by the fact that there may in fact be some relevant information in the temporal coherence derived from the consistency of the phase over time. Because vegetation tends to result in lower temporal coherence, this could be an additional feature for land cover classification.

The main problem, however, remains the lack of reliable, easily accessible reference data. Deforestation detection essentially constitutes a classification problem: it can be viewed e.g. as classifying each pixel as forest vs. non-forest or change vs. no change. Any classification algorithm that is not purely based on prior knowledge, i.e. the physics of underlying processes generating the data we observe, relies heavily on the availability and quality of reference data that can be used to train and assess these algorithms. Both the availability and quality of reference data for deforestation monitoring are lacking. Where such data are available at desirable temporal and spatial scales they are generally the output of a classification algorithm (e.g. Hansen, Potapov, et al. (2013)) and not, therefore, strictly ground truth data. Actual ground data are very scarce and often cover limited spatial extents with infrequent revisits, if any. Furthermore, as such data are collected in single forest biomes they generally do not allow for generalization to global scales. In reality, reference data are often derived from manual/visual inspection of high-resolution satellite imagery, e.g. Google Earth or commercial data (e.g., Planet). As commercial high-resolution imagery is expensive and the inspection process labor intensive, these data are often not publicly available.

For this reason, I mainly focused on testing and developing change detection algorithms that either do not require reference data or are robust to errors in the reference data, i.e. noisy training labels. As an example of such a change detection algorithm I chose to implement the Omnibus test introduced by Conradsen, Nielsen, and Skriver (2016). In addition to not requiring any reference data, this algorithm is making full use of the dual-polarimetric Sentinel-1 data by deriving test statistics from the complex Wishart distribution (a generalization of normal distributions for complex covariance matrices). The implementation of this polarimetric change detection algorithm as discussed in section 4.4 has shown that Sentinel-1 data are indeed capable of picking up deforestation. For the selected area there was an omission error of 39.3% and a commission error of 2.5% compared to the reference data. It would thus be a mistake to disregard Sentinel-1 data for this purpose, especially as the open data policy and frequent global coverage will certainly prove valuable for forest monitoring. The issue with any change detection algorithm that does not rely on reference data like the Omnibus test is that it will pick up any type of significant change indiscriminately, whether this change corresponds to deforestation, seasonal changes, agricultural harvest and growth cycles, or other artificial or natural disturbance events.

I went on to investigate the information content in the spatial context, and how this would improve the change detection results obtained from the Omnibus test algorithm. I introduced data cubes as an intuitive way to represent multivariate spatio-temporal data. Spatio-temporal filters were introduced and assessed as a simple way to generate features that capture spatial context. There was a clear improvement of visual image quality when applying edge-preserving filters leading also to improved change detection results. Ultimately, these results clearly indicated that spatial context as well as the temporal signature *do* contain useful information for deforestation detection and as such should not be disregarded.

Based on the limitations of change detection algorithms that are completely agnostic to the type of change, and the lack of reliable reference data, I followed a trade-off approach and developed a method that was able to address both of these challenges. The basic strategy for this method was detecting changes as deviations from a reference class such as forest, where the reference class is characterized by a reference map that can potentially be of low quality. Using this method, a near-perfect change detection sensitivity was achieved (producer's accuracy above 99%), although false positives resulted in a low user's accuracy of about 60%. The false positives mainly occurred in areas that were non-forest to begin with and should be easily filtered. The mean change detection delay based on visual interpretation of Landsat images amounts to about two to three months. Additional efforts are required to reduce this detection delay by calibrating the change threshold appropriately. The proposed change detection algorithm has been shown to be robust to noisy reference data. It still delivers reliable change detection results without significant degradation with up to 10% of the reference forest mask mislabelled. It is a novel addition to the existing arsenal of change detection algorithms, as it fills a gap between the fully supervised and fully unsupervised methods.

The quality of the underlying reference data must be considered when interpreting and comparing the accuracy scores presented throughout the various results in this thesis. None of the data used for testing and validation were in fact strictly ground truth data, as they were all obtained from either visual interpretation or machine learning based on optical medium to high-resolution aerial or satellite data. As such, the reference data may be subject to their own biases and errors and the true errors of both commission and omission for the change detection algorithms discussed in this thesis may in fact be higher or lower than stated.

## 7.1 Future Work

There is an open research question regarding the relevance of phase information for forest/non-forest discrimination. Preliminary results have indicated that with some exceptions, there is barely any useful information in the VV-VH phase difference when it comes to separating forest from non-forest pixels. In order to discard the phase entirely as a useful quantity when looking at land cover classification, additional research needs to be done regarding derived interferometric quantities, in particular the temporal coherence. This has only recently been addressed by Borlaf-Mena et al. (2021), who found that the overall accuracy may improve by up to 5% when adding temporal coherence as a feature.

More work is needed regarding the inclusion of spatial context into the change detection process. The previous results can be extended by a more detailed analysis of the exact features that capture the most difference between forest and non-forest. Furthermore, there is a possibility to develop an entirely different change detection algorithm that incorporates spatial context directly and thus moves beyond pixelwise analysis, without relying on spatial features. Such an algorithm could be implemented as a multi-dimensional edge detection. This would mean to abandon the concept of change detection as a temporal phenomenon and acknowledging that space and time dimensions are fundamentally equivalent. Accurately delineating forest boundaries as surfaces in space-time would directly solve the problem of deforestation detection.

While the preliminary results from the change detection algorithm based on temporal reference class deviation are very promising, several adjustments can still be made to improve its performance. Firstly, the temporal accuracy of the detected change should be optimized by carefully calibrating the change point extraction from the cumulative distance curve. Secondly, some of the false positives that are causing a low UA over non-forest areas could be eliminated by adding a statistical test for the significance of a detected change. This would in particular remove erroneous changes near the beginning or the end of the time series where not enough data points are available before or after the detected change for a statistically significant change. Finally, a more stringent validation of the method is required, including testing over a wide range of forest types.

Additionally, the method needs to be formally extended to include multi-sensor data, in particular optical data such as Landsat or Sentinel-2. This step comes with challenges as discussed in section 4.6 including deviating spatial and temporal resolutions, as well as an increased amount of missing data for optical sensors due to cloud cover. However, when treated as a data science problem, the data set will not get much more complex since SAR data are already multivariate and the addition of optical data would merely increase the number of variables available at each pixel. There is additional potential to more deeply explore polarimetry and interferometry in the context of forest monitoring. One variable of particular interest is the temporal coherence and its preservation over forested vs. deforested area.

In the process of testing change detection algorithms on a study area containing both agriculture and forest, it has emerged that the change signature may in fact be indicative of the type of land cover and even land use, as different forms of vegetation and crop types undergo deviating seasonal patterns due to growth and harvest cycles. Another research question may thus address the aptitude of change signatures to classify land cover types.

While not the main focus of my PhD research, the development of a software framework as described in chapter 6 has become a natural necessity for the implementation of any of the algorithms discussed. With every new data source and every new algorithm implemented, the library has already grown and evolved naturally, if only to facilitate my own work. My hope is that other researchers in the Earth observation community will find some of the tools useful and that it will contribute to a more unified remote sensing ecosystem in Python.

# Chapter 8

# Conclusions

I established that Sentinel-1 data are capable of reliably distinguishing forest and non-forest for a variety of different forest types globally. My study of six diverse forest sites in Alaska, Colombia, Finland, Florida, Indonesia, and the UK demonstrated an accuracy of about 90% forest identification using the annual mean and standard deviation of co- and cross-polarized backscatter. While optical data may still be the superior option for generating land cover maps at low temporal repeat intervals, SAR has the unique advantage of being independent of cloud cover and is thus the only choice for a near-real-time monitoring system with complete coverage.

After testing the performance of forest/non-forest classification and reviewing and implementing several change detection algorithms, it was clear that the quality of all results was highly dependent on the quality of the input, i.e. the reference or training data. With the lack of reliable reference data being one of the most common challenges in designing change detection algorithms, this was limiting the overall usefulness of these methods for the general case. On the other hand, unsupervised methods like the Omnibus test are good at picking up changes in the data but exhibit a low specificity to deforestation events. They are therefore not in general suitable for forest monitoring. As a consequence of these findings I converged on a semi-supervised approach to deforestation-detection.

While traditional binary classification algorithms decide class membership based on the similarity to both classes, the method I developed is assigning a measure of forest similarity based only on how well the pixel is represented by the reference forest class. Because it only requires a rough map of the reference class (forest), this approach is more robust and more readily generalizable than other methods. In addition, the nature and composition of non-forest areas varies vastly between different regions, and attempting to capture these in one homogeneous class is sure to introduce overfitting. My method does not attempt to parametrize the non-forest class and therefore only requires an adaptation of the forest model for scaling up globally. Because of its overall simplicity I believe this method has the potential to be further improved and to be widely adopted for change detection.

Circling back to the research goal formulated in the introduction, the proposed method fulfills the requirements that I set out to achieve:

1. It is agnostic to the physics of the underlying processes, as all prior information about forest behavior is encoded in the reference forest class.
2. It is robust in the presence of sub-optimal reference data.
3. The method scales linearly with space and is trivially parallelizable. The training process only consists of computing the forest time series ensemble, which is extremely fast compared to most machine learning model fitting.

There is always a trade-off between spatial and temporal accuracy when detecting changes. Much of the decision making around algorithm parameters is therefore based around striking a suitable balance between spatial resolution and detection timeliness. Depending on the specific application, either one may be the more important target.

With the anticipated launch of a range of spaceborne instruments with various new sensors there is an exciting opportunity for the development of advanced algorithms for multi-sensor fusion. Ultimately, the amount and quality of available data is increasing at a rapid pace. As long as our algorithms can keep up I am hopeful that satellite based Earth observation will continue to be a force for good and a key strategic instrument in monitoring and mitigating the deliberate and inadvertent destruction of our planet and the effects of climate change more generally.

# Bibliography

Adams, J. B., Sabol, D. E., Kapos, V., Filho, R. A., Roberts, D., Smith, M. O., & Gillespie, A. R. (1995). Classification of multispectral images based on fractions and endmembers: applications to land-cover change in the Brazilian Amazon. *Remote Sensing of Environment*, *52*(94), 137–152. https://doi.org/10.1016/0034-4257(94)00098-8

Aleksandrov, M., Batič, M., Milčinski, G., Korais, D., Kadunc, M., Lubej, M., Lukšič, Ž., Peressutti, D., Slijepčević, T., Vesel, N., Višnjić, J., Zupanc, A., Puc, J., Žust, L., Sovdat, B., Burja, A., Erzin, E., Fournier, H., Koprivec, F., . . . Bollinger, D. (2022). Eo-learn [Accessed: 2022-06-25]. https://github.com/sentinel-hub/eo-learn

Anfinsen, S. N., Member, S., Doulgeris, A. P., Member, S., & Eltoft, T. (2009). Estimation of the Equivalent Number of Looks in Polarimetric Synthetic Aperture Radar Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *47*(11), 3795–3809. https://doi.org/10.1109/TGRS.2009.2019269

Askne, J., & Smith, G. (1997). Forest insar decorrelation and classification properties. *ERS SAR Interferometry*, *406*, 95.

Askne, J. I. H., Dammert, P. B. G., Ulander, L. M. H., & Smith, G. (1997). C-band repeat-pass interferometric SAR observations of the forest. *IEEE Transactions on Geoscience and Remote Sensing*, *35*(1), 25–35.

Balzter, H. (2001). Forest mapping and monitoring with interferometric synthetic aperture radar (InSAR). *Progress in Physical Geography*, *25*(2), 159–177. https://doi.org/10.1191/030913301666986397

Balzter, H., Luckman, A., Skinner, L., Rowland, C., & Dawson, T. (2007). Observations of forest stand top height and mean height from interferometric SAR and LiDAR over a conifer plantation at Thetford Forest, UK. *International Journal of Remote Sensing*, *28*(6), 1173–1197. https://doi.org/10.1080/01431160600904998

Balzter, H., Rowland, C. S., & Saich, P. (2007). Forest canopy height and carbon estimation at Monks Wood National Nature Reserve, UK, using dual-wavelength SAR interferometry. *Remote Sensing of Environment*, *108*(3), 224–239. https://doi.org/https://doi.org/10.1016/j.rse.2006.11.014

Balzter, H., Cole, B., Thiel, C., & Schmullius, C. (2015). Mapping CORINE Land Cover from Sentinel-1A SAR and SRTM Digital Elevation Model Data using Random Forests. *Remote Sensing*, *7*(11), 14876–14898. https://doi.org/10.3390/rs71114876

Balzter, H., Talmon, E., Wagner, W., Gaveau, D., Plummer, S., Yu, J. J., Quegan, S., Davidson, M., Toan, T. L., Gluck, M., Shvidenko, A., Nilsson, S., Tansey, K., Luckman, A., & Schmullius, C. (2002). Accuracy assessment of a large-scale forest cover map of central siberia from synthetic aperture radar. *Canadian Journal of Remote Sensing*, *28*(6), 719–737. https://doi.org/10.5589/m02-067

Baret, F., & Guyot, G. (1991). Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sensing of Environment*, *35*(2-3), 161–173. https://doi.org/10.1016/0034-4257(91)90009-U

Baret, F., Guyot, G., & Major, D. J. (1989). TSAVI: A Vegetation Index Which Minimizes Soil Brightness Effects On LAI And APAR Estimation. *12th Canadian Symposium on Remote Sensing Geoscience and Remote Sensing Symposium, 3*(2), 1355–1358. https://doi.org/10.1109/IGARSS.1989.576128

Borlaf-Mena, I., Badea, O., & Tanase, M. A. (2021). Assessing the Utility of Sentinel-1 Coherence Time Series for Temperate and Tropical Forest Mapping. *Remote Sensing 2021, Vol. 13, Page 4814*, *13*(23), 4814. https://doi.org/10.3390/RS13234814

Böttcher, H., Eisbrenner, K., Fritz, S., Kindermann, G., Kraxner, F., McCallum, I., & Obersteiner, M. (2009). An assessment of monitoring requirements and costs of 'Reduced Emissions from Deforestation and Degradation'. *Carbon Balance and Management*, *4*, 7. https://doi.org/10.1186/1750-0680-4-7

Bouvet, A., Mermoz, S., Ballère, M., Koleck, T., & Le Toan, T. (2018). Use of the SAR shadowing effect for deforestation detection with Sentinel-1 time series. *Remote Sensing*, *10*(8), 1–19. https://doi.org/10.3390/rs10081250

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Bringmann, K., & Künnemann, M. (2015). Quadratic conditional lower bounds for string problems and dynamic time warping. *2015 IEEE 56th Annual Sym-*

*posium on Foundations of Computer Science*, 79–97. https://doi.org/10.1109/FOCS.2015.15

Brodersen, K. H., Ong, C. S., Stephan, K. E., & Buhmann, J. M. (2010). The balanced accuracy and its posterior distribution. *Proceedings - International Conference on Pattern Recognition*, 3121–3124. https://doi.org/10.1109/ICPR.2010.764

Buades, A., Coll, B., & Morel, J.-M. (2011). Non-Local Means Denoising. *Image Processing On Line*, *1*, 208–212. https://doi.org/10.5201/ipol.2011.bcm_nlm

Bunting, P., Clewley, D., Lucas, R. M., & Gillingham, S. (2014). The remote sensing and gis software library (rsgislib). *Computers & Geosciences*, *62*, 216–226. https://doi.org/https://doi.org/10.1016/j.cageo.2013.08.007

Buschmann, C., & Nagel, E. (1993). In vivo spectroscopy and internal optics of leaves as basis for remote sensing of vegetation. *International Journal of Remote Sensing*, *14*(4), 711–722. https://doi.org/10.1080/01431169308904370

Calders, K., Adams, J., Armston, J., Bartholomeus, H., Bauwens, S., Bentley, L. P., Chave, J., Danson, F. M., Demol, M., Disney, M., Gaulton, R., Krishna Moorthy, S. M., Levick, S. R., Saarinen, N., Schaaf, C., Stovall, A., Terryn, L., Wilkes, P., & Verbeeck, H. (2020). Terrestrial laser scanning in forest ecology: Expanding the horizon. *Remote Sensing of Environment*, *251*(September), 112102. https://doi.org/10.1016/j.rse.2020.112102

Calders, K., Newnham, G., Burt, A., Murphy, S., Raumonen, P., Herold, M., Culvenor, D., Avitabile, V., Disney, M., Armston, J., & Kaasalainen, M. (2015). Nondestructive estimates of above-ground biomass using terrestrial laser scanning. *Methods in Ecology and Evolution*, *6*(2), 198–208. https://doi.org/10.1111/2041-210X.12301

Calero, E. J. (2018). *Sentinels POD Service File Format Specifications* (tech. rep.). ESA. Frascati. https://sentinel.esa.int/documents/247904/351187/GMES%7B%5C_%7DSentinels%7B%5C_%7DPOD%7B%5C_%7DService%7B%5C_%7DFile%7B%5C_%7DFormat%7B%5C_%7DSpecification

Cartus, O., Santoro, M., Wegmuller, U., Labriere, N., & Chave, J. (2021). Sentinel-1 Coherence for Mapping Above-Ground Biomass in Semiarid Forest Areas. *IEEE Geoscience and Remote Sensing Letters*, 1–5. https://doi.org/10.1109/LGRS.2021.3071949

Chave, J., Andalo, C., Brown, S., Cairns, M. A., Chambers, J. Q., Eamus, D., Fölster, H., Fromard, F., Higuchi, N., Kira, T., Lescure, J. P., Nelson, B. W., Ogawa, H., Puig, H., Riéra, B., & Yamakura, T. (2005). Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia*, *145*(1), 87–99. https://doi.org/10.1007/s00442-005-0100-x

Chen, L., Ren, C., Zhang, B., Wang, Z., & Xi, Y. (2018). Estimation of forest above-ground biomass by geographically weighted regression and machine learning with sentinel imagery. *Forests*, *9*(10), 1–20. https://doi.org/10.3390/f9100582

Choi, H. R., & Kim, T. Y. (2017). Combined Dynamic Time Warping with Multiple Sensors for 3D Gesture Recognition. *Sensors (Switzerland)*, *17*(8). https://doi.org/10.3390/s17081893

Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A Seasonal-Trend Decomposition Procedure Based on Loess. *Journal of Official Statistics*, *6*(1), 3–73.

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Cohen, W. B., Yang, Z., & Kennedy, R. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync - Tools for calibration and validation. *Remote Sensing of Environment*, *114*(12), 2911–2924. https://doi.org/10.1016/j.rse.2010.07.010

Comer, D., Chapman, B., & Comer, J. (2017). Detecting Landscape Disturbance at the Nasca Lines Using SAR Data Collected from Airborne and Satellite Platforms. *Geosciences*, *7*(4), 106. https://doi.org/10.3390/geosciences7040106

Conradsen, K., Nielsen, A., Schou, J., & Skriver, H. (2003). A test statistic in the complex wishart distribution and its application to change detection in polarimetric SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(1), 4–19. https://doi.org/10.1109/TGRS.2002.808066

Conradsen, K., Nielsen, A. A., & Skriver, H. (2016). Determining the Points of Change in Time Series of Polarimetric SAR Data. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(5), 3007–3024. https://doi.org/10.1109/TGRS.2015.2510160

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. https://doi.org/10.1007/BF00994018

Crist, E. P. (1985). A TM Tasseled Cap equivalent transformation for reflectance factor data. *Remote Sensing of Environment*, *17*(3), 301–306. https://doi.org/10.1016/0034-4257(85)90102-6

Crist, E. P., & Cicone, R. C. (1984). A Physically-Based Transformation of Thematic Mapper Data - The Tasseled Cap. *IEEE Transactions on Geoscience and*

*Remote Sensing*, *22*(3), 256–263. https://doi.org/10.1109/TGRS.1984.350619

De Zan, F., & Guarnieri, A. M. (2006). TOPSAR: Terrain Observation by Progressive Scans. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(9), 2352–2360. https://doi.org/10.1109/TGRS.2006.873853

Deutscher, J., Perko, R., Gutjahr, K., Hirschmugl, M., & Schardt, M. (2013). Mapping Tropical Rainforest Canopy Disturbances in 3D by COSMO-SkyMed Spotlight InSAR-Stereo Data to Detect Areas of Forest Degradation. *Remote Sensing*, *5*(2), 648–663. https://doi.org/10.3390/rs5020648

DeVries, B., Verbesselt, J., Kooistra, L., & Herold, M. (2015). Robust monitoring of small-scale forest disturbances in a tropical montane forest using Landsat time series. *Remote Sensing of Environment*, *161*, 107–121. https://doi.org/10.1016/j.rse.2015.02.012

Dobson, M. C., Ulaby, F. T., Le Toan, T., Beaudoin, A., Kasischke, E. S., & Christensen, N. (1992). Dependence of Radar Backscatter on Coniferous Forest Biomass. *IEEE Transactions on Geoscience and Remote Sensing*, *30*(2), 412–415. https://doi.org/10.1109/36.134090

Dostálová, A., Hollaus, M., Milenković, M., & Wagner, W. (2016). Forest Area Derivation From Sentinel-1 Data. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *III-7*(July), 227–233. https://doi.org/10.5194/isprsannals-iii-7-227-2016

Dutrieux, L. P., Verbesselt, J., Kooistra, L., & Herold, M. (2015). Monitoring forest cover loss using multiple data streams, a case study of a tropical dry forest in Bolivia. *ISPRS Journal of Photogrammetry and Remote Sensing*, *107*, 112–125. https://doi.org/10.1016/j.isprsjprs.2015.03.015

Felzenszwalb, P. F., & Huttenlocher, D. P. (2004). Efficient graph-based image segmentation. *International Journal of Computer Vision*, *59*(2), 167–181. https://doi.org/10.1023/B:VISI.0000022288.19776.77

Forkuor, G., Benewinde Zoungrana, J. B., Dimobe, K., Ouattara, B., Vadrevu, K. P., & Tondoh, J. E. (2020). Above-ground biomass mapping in West African dryland forest using Sentinel-1 and 2 datasets - A case study. *Remote Sensing of Environment*, *236*(November 2019), 111496. https://doi.org/10.1016/j.rse.2019.111496

García, M., Saatchi, S., Ustin, S., & Balzter, H. (2018). Modelling forest canopy height by integrating airborne LiDAR samples with satellite Radar and multispectral imagery. *International Journal of Applied Earth Observation and*

*Geoinformation*, *66*(November 2017), 159–173. https://doi.org/10.1016/j.jag.2017.11.017

Gaveau, D. L. A., Balzter, H., & Plummer, S. (2003). Forest woody biomass classification with satellite-based radar coherence over 900 000 km2 in Central Siberia. *Forest Ecology and Management*, *174*(1), 65–75. https://doi.org/https://doi.org/10.1016/S0378-1127(02)00028-2

GDAL Development Team. (2018). *GDAL - Geospatial Data Abstraction Library, Version 2.3.2*. Open Source Geospatial Foundation.

Gitelson, A. A., Kaufman, Y. J., & Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment*, *58*(3), 289–298. https://doi.org/10.1016/S0034-4257(96)00072-7

Gitelson, A. A., Kaufman, Y. J., Stark, R., & Rundquist, D. (2002). Novel algorithms for remote estimation of vegetation fraction. *Remote Sensing of Environment*, *80*(1), 76–87. https://doi.org/10.1016/S0034-4257(01)00289-9

Gold, O., & Sharir, M. (2018). Dynamic Time Warping and Geometric Edit Distance: Breaking the Quadratic Barrier. *ACM Trans. Algorithms*, *14*(4). https://doi.org/10.1145/3230734

Gonzalez de Tanago, J., Lau, A., Bartholomeus, H., Herold, M., Avitabile, V., Raumonen, P., Martius, C., Goodman, R. C., Disney, M., Manuri, S., Burt, A., & Calders, K. (2018). Estimation of above-ground biomass of large tropical trees with terrestrial LiDAR. *Methods in Ecology and Evolution*, *9*(2), 223–234. https://doi.org/10.1111/2041-210X.12904

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*. https://doi.org/10.1016/j.rse.2017.06.031

Hamunyela, E., Reiche, J., Verbesselt, J., & Herold, M. (2017). Using space-time features to improve detection of forest disturbances from Landsat time series. *Remote Sensing*, *9*(6), 1–17. https://doi.org/10.3390/rs9060515

Hamunyela, E., Verbesselt, J., & Herold, M. (2016). Using spatial context to improve early detection of deforestation from Landsat time series. *Remote Sensing of Environment*, *172*, 126–138. https://doi.org/10.1016/j.rse.2015.11.006

Hancock, S., McGrath, C., Lowe, C., Davenport, I., & Woodhouse, I. (2021). Requirements for a global lidar system: Spaceborne lidar with wall-to-wall coverage. *Royal Society Open Science*.

Hansen, J. (2021). *Jnhansen/nd: V0.3.1* (Version 0.3.1). Zenodo. https://doi.org/10.5281/zenodo.5637053

Hansen, J. N. (2022). Nd – a framework for the analysis of n-dimensional earth observation data. *Journal of Open Research Software*, *10*(1), 3. https://doi.org/10.5334/jors.377

Hansen, J. N., Mitchard, E. T. A., & King, S. (2020). Assessing Forest/Non-Forest Separability Using Sentinel-1 C-Band Synthetic Aperture Radar. *Remote Sensing*, *12*(11). https://doi.org/10.3390/rs12111899

Hansen, M. C., DeFries, R. S., Townshend, J. R. G., Carroll, M., Dimiceli, C., & Sohlberg, R. A. (2003). Global Percent Tree Cover at a Spatial Resolution of 500 Meters: First Results of the MODIS Vegetation Continuous Fields Algorithm. *Earth Interactions*, *7*(10), 1–15. https://doi.org/10.1175/1087-3562(2003)007⟨0001:GPTCAA⟩2.0.CO;2

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., & Townshend, J. R. G. (2013). High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science*, *342*(November), 850–854. https://doi.org/10.1126/science.1244693

Healey, S. P., Cohen, W. B., Zhiqiang, Y., & Krankina, O. N. (2005). Comparison of Tasseled Cap-based Landsat data structures for use in forest disturbance detection. *Remote Sensing of Environment*, *97*(3), 301–310. https://doi.org/10.1016/j.rse.2005.05.009

Herold, M., Román-Cuesta, R. M., Mollicone, D., Hirata, Y., Van Laake, P., Asner, G. P., Souza, C., Skutsch, M., Avitabile, V., & MacDicken, K. (2011). Options for monitoring and estimating historical carbon emissions from forest degradation in the context of REDD+. *Carbon Balance and Management*, *6*(13), 1–7. https://doi.org/10.1186/1750-0680-6-13

Herold, M., & Skutsch, M. (2011). Monitoring, reporting and verification for national REDD + programmes: Two proposals. *Environmental Research Letters*, *6*(1). https://doi.org/10.1088/1748-9326/6/1/014002

Hoekman, D. H., Vissers, M. A., & Wielaard, N. (2010). PALSAR Wide-Area Mapping of Borneo: Methodology and Map Validation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *3*(4), 605–617. https://doi.org/10.1109/JSTARS.2010.2070059

Hosonuma, N., Herold, M., De Sy, V., De Fries, R. S., Brockhaus, M., Verchot, L., Angelsen, A., & Romijn, E. (2012). An assessment of deforestation and forest degradation drivers in developing countries. *Environmental Research Letters*, *7*(4). https://doi.org/10.1088/1748-9326/7/4/044009

Hoyer, S., & Hamman, J. J. (2017). xarray: N-D labeled Arrays and Datasets in Python. *Journal of Open Research Software*, *5*. https://doi.org/10.5334/jors.148

Huang, C., Goward, S. N., Masek, J. G., Thomas, N., Zhu, Z., & Vogelmann, J. E. (2010). An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sensing of Environment*, *114*(1), 183–198. https://doi.org/10.1016/j.rse.2009.08.017

Huete, A., Justice, C., & Liu, H. (1994). Development of vegetation and soil indices for MODIS-EOS. *Remote Sensing of Environment*, *49*(3), 224–234. https://doi.org/10.1016/0034-4257(94)90018-3

Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, *25*(3), 295–309. https://doi.org/10.1016/0034-4257(88)90106-X

INPE. (n.d.). DEGRAD - Mapeamento da Degradação Florestal na Amazônia Brasileira. Retrieved May 21, 2018, from http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/degrad

IPCC. (2003). *Good practice guidance for land use, land use change and forestry* (tech. rep.). https://doi.org/citeulike-article-id:1260638

Jin, H., & Eklundh, L. (2014). A physically based vegetation index for improved monitoring of plant phenology. *Remote Sensing of Environment*, *152*, 512–525. https://doi.org/10.1016/J.RSE.2014.07.010

Joshi, N., Baumann, M., Ehammer, A., Fensholt, R., Grogan, K., Hostert, P., Jepsen, M. R., Kuemmerle, T., Meyfroidt, P., Mitchard, E. T., Reiche, J., Ryan, C. M., & Waske, B. (2016). A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring. https://doi.org/10.3390/rs8010070

Joshi, N., Mitchard, E. T., Woo, N., Torres, J., Moll-Rocek, J., Ehammer, A., Collins, M., Jepsen, M. R., & Fensholt, R. (2015). Mapping dynamics of deforestation and forest degradation in tropical forests using radar satellite data. *Environmental Research Letters*, *10*(3), 34014. https://doi.org/10.1088/1748-9326/10/3/034014

JPL OurOcean. (2010). GHRSST Level 4 G1SST Global Foundation Sea Surface Temperature Analysis. *NASA Physical Oceanography DAAC*. https://doi.org/10.5067/GHG1S-4FP01

Juang, B.-H. (1984). On the Hidden Markov Model and Dynamic Time Warping for Speech Recognition — A Unified View. *AT T Bell Laboratories Technical Journal*, *63*(7), 1213–1243. https://doi.org/10.1002/j.1538-7305.1984.tb00034.x

Kaasalainen, S., Holopainen, M., Karjalainen, M., Vastaranta, M., Kankare, V., Karila, K., & Osmanoglu, B. (2015). Combining lidar and Synthetic Aperture Radar data to estimate forest biomass: Status and prospects. *Forests*, *6*(1), 252–270. https://doi.org/10.3390/f6010252

Kennedy, R. E., Cohen, W. B., & Schroeder, T. A. (2007). Trajectory-based change detection for automated characterization of forest disturbance dynamics. *Remote Sensing of Environment*, *110*(3), 370–386. https://doi.org/10.1016/j.rse.2007.03.010

Kennedy, R. E., Yang, Z., & 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*(12), 2897–2910. https://doi.org/10.1016/j.rse.2010.07.008

Killick, R., Fearnhead, P., & Eckley, I. A. (2012). Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association*, *107*(500), 1590–1598. https://doi.org/10.1080/01621459.2012.737745

Koenig, L. L., Lucero, J. C., & Perlman, E. (2008). Speech production variability in fricatives of children and adults: Results of functional data analysis. *The Journal of the Acoustical Society of America*, *124*(5), 3158–3170. https://doi.org/10.1121/1.2981639

Kwak, D. A., Lee, W. K., Lee, J. H., Biging, G. S., & Gong, P. (2007). Detection of individual trees and estimation of tree height using LiDAR data. *Journal of Forest Research*, *12*(6), 425–434. https://doi.org/10.1007/s10310-007-0041-9

Laliberte, A. S., Browning, D. M., & Rango, A. (2012). A comparison of three feature selection methods for object-based classification of sub-decimeter resolution UltraCam-L imagery. *International Journal of Applied Earth Observation and Geoinformation*, *15*(1), 70–78. https://doi.org/10.1016/j.jag.2011.05.011

Laporte, N. T., Stabach, J. A., Grosch, R., Lin, T. S., & Goetz, S. J. (2007). Expansion of industrial logging in Central Africa. *Science*, *316*(5830), 1451. https://doi.org/10.1126/science.1141057

Laurin, G. V., Balling, J., Corona, P., Mattioli, W., Papale, D., Puletti, N., Rizzo, M., Truckenbrodt, J., & Urban, M. (2018). Above-ground biomass prediction by Sentinel-1 multitemporal data in central Italy with integration of ALOS2 and Sentinel-2 data. *Journal of Applied Remote Sensing*, *12*(01), 1. https://doi.org/10.1117/1.jrs.12.016008

Lee, J.-S. (1981). Refined Filtering of Image Noise Using Local Statistics. *Computer Graphics and Image Processing*, *15*(4), 380–389. https://doi.org/10.1016/S0146-664X(81)80018-4

Lee, J.-S. (1983). Digital image smoothing and the sigma filter. *Computer Vision, Graphics and Image Processing*, *24*(2), 255–269. https://doi.org/10.1016/0734-189X(83)90047-6

Lee, J.-S., & Pottier, E. (2009). *Polarimetric Radar Imaging: From Basics to Applications*. https://doi.org/10.1201/9781420054989.fmatt

Lee, J.-S., Wen, J.-H., Ainsworth, T. L., Chen, K.-S., & Chen, A. J. (2009). Improved Sigma Filter for Speckle Filtering of SAR Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *47*(1), 202–213. https://doi.org/10.1109/TGRS.2008.2002881

Lehmann, E. A., Caccetta, P., Lowell, K., Mitchell, A., Zhou, Z. S., Held, A., Milne, T., & Tapley, I. (2015). SAR and optical remote sensing: Assessment of complementarity and interoperability in the context of a large-scale operational forest monitoring system. *Remote Sensing of Environment*, *156*, 335–348. https://doi.org/10.1016/j.rse.2014.09.034

Li, Y., Li, M., Li, C., & Liu, Z. (2020). Forest aboveground biomass estimation using Landsat 8 and Sentinel-1A data with machine learning algorithms. *Scientific Reports*, *10*(1), 1–12. https://doi.org/10.1038/s41598-020-67024-3

Lindquist, E. J., D'Annunzio, R., Gerrand, A., MacDicken, K., Achard, F., Beuchle, R., Brink, A., Eva, H. D., Mayaux, P., San-Miguel-Ayanz, J., & Stibig, H.-J. (2012). *Global forest land-use change 1990-2005* (tech. rep.). FAO & JRC.

Lu, D. (2006). The potential and challenge of remote sensing-based biomass estimation. *International Journal of Remote Sensing*, *27*(7), 1297–1328. https://doi.org/10.1080/01431160500486732

Lucas, R. M., Milne, A. K., Cronin, N., Witte, C., & Denham, R. (2000). The potential of synthetic aperture radar (sar) for quantifying the biomass of australia's woodlands. *Rangeland Journal*, *22*(1), 124–140. https://doi.org/10.1071/RJ0000124

Lucas, R., Bunting, P., Clewley, D., Armston, J., Fairfax, R., Fensham, R., Accad, A., Kelley, J., Laidlaw, M., Eyre, T., Bowen, M., Carreiras, J., Bray, S., Metcalfe, D., Dwyer, J., & Shimada, M. (2010). An Evaluation of the ALOS PALSAR L-Band Backscatter—Above Ground Biomass Relationship Queensland, Australia: Impacts of Surface Moisture Condition and Vegetation Structure. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *3*(4), 576–593. https://doi.org/10.1109/JSTARS.2010.2086436

Luckman, A., Baker, J., Honzák, M., & Lucas, R. (1998). Tropical forest biomass density estimation using JERS-1 SAR: Seasonal variation, confidence limits, and application to image mosaics. *Remote Sensing of Environment*, *63*(2), 126–139. https://doi.org/10.1016/S0034-4257(97)00133-8

Masek, J. G., Huang, C., Wolfe, R., Cohen, W., Hall, F., Kutler, J., & Nelson, P. (2008). North American forest disturbance mapped from a decadal Landsat record. *Remote Sensing of Environment*, *112*(6), 2914–2926. https://doi.org/10.1016/j.rse.2008.02.010

Mayaux, P., Eva, H., Gallego, J., Strahler, A. H., Herold, M., Agrawal, S., Naumov, S., De Miranda, E. E., Di Bella, C. M., Ordoyne, C., Kopin, Y., & Roy, P. S. (2006). Validation of the Global Land Cover 2000 Map. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(7), 1728–1737. https://doi.org/10.1109/TGRS.2006.864370

McRoberts, R. E., & Tomppo, E. O. (2007). Remote sensing support for national forest inventories. *Remote Sensing of Environment*, *110*(4), 412–419. https://doi.org/10.1016/j.rse.2006.09.034

Mermoz, S., & Le Toan, T. (2016). Forest disturbances and regrowth assessment using ALOS PALSAR data from 2007 to 2010 in Vietnam, Cambodia and Lao PDR. *Remote Sensing*, *8*(3), 1–22. https://doi.org/10.3390/rs8030217

Mermoz, S., Le Toan, T., Villard, L., Réjou-Méchain, M., & Seifert-Granzin, J. (2014). Biomass assessment in the Cameroon savanna using ALOS PALSAR data. *Remote Sensing of Environment*, *155*, 109–119. https://doi.org/10.1016/j.rse.2014.01.029

Met Office. (2010 - 2020). *Cartopy: A cartographic python library with a matplotlib interface*. Exeter, Devon. http://scitools.org.uk/cartopy

Mitchard, E. T. A., Saatchi, S. S., Lewis, S. L., Feldpausch, T. R., Woodhouse, I. H., Sonké, B., Rowland, C., & Meir, P. (2011). Measuring biomass changes due to woody encroachment and deforestation/degradation in a forest-savanna boundary region of central Africa using multi-temporal L-band radar backscatter. *Remote Sensing of Environment*, *115*(11), 2861–2873. https://doi.org/10.1016/j.rse.2010.02.022

Mitchard, E. T. A., Saatchi, S. S., White, L. J. T., Abernethy, K. A., Jeffery, K. J., Lewis, S. L., Collins, M., Lefsky, M. A., Leal, M. E., Woodhouse, I. H., & Meir, P. (2012). Mapping tropical forest biomass with radar and spaceborne LiDAR in Lopé National Park, Gabon: Overcoming problems of high biomass and persistent cloud. *Biogeosciences*, *9*(1), 179–191. https://doi.org/10.5194/bg-9-179-2012

Mitchard, E. T. A., Saatchi, S. S., Woodhouse, I. H., Nangendo, G., Ribeiro, N. S., Williams, M., Ryan, C. M., Lewis, S. L., Feldpausch, T. R., & Meir, P. (2009). Using satellite radar backscatter to predict above-ground woody biomass: A consistent relationship across four different African landscapes. *Geophysical Research Letters*, *36*(23), 1–6. https://doi.org/10.1029/2009GL040692

Mitchard, E. T. A., Saatchi, S. S., Baccini, A., Asner, G. P., Goetz, S. J., Harris, N. L., & Brown, S. (2013). Uncertainty in the spatial distribution of tropical forest biomass: A comparison of pan-tropical maps. *Carbon Balance and Management*, *8*(1), 1–13. https://doi.org/10.1186/1750-0680-8-10

Mitchard, E. T., Feldpausch, T. R., Brienen, R. J., Lopez-Gonzalez, G., Monteagudo, A., Baker, T. R., Lewis, S. L., Lloyd, J., Quesada, C. A., Gloor, M., ter Steege, H., Meir, P., Alvarez, E., Araujo-Murakami, A., Aragão, L. E., Arroyo, L., Aymard, G., Banki, O., Bonal, D., . . . Phillips, O. L. (2014). Markedly divergent estimates of Amazon forest carbon density from ground plots and satellites. *Global Ecology and Biogeography*, *23*(8), 935–946. https://doi.org/10.1111/geb.12168

Mitchell, A. L., Rosenqvist, A., & Mora, B. (2017). Current remote sensing approaches to monitoring forest degradation in support of countries measurement, reporting and verification (MRV) systems for REDD+. *Carbon Balance and Management*, *12*(1), 9. https://doi.org/10.1186/s13021-017-0078-9

Mitchell, A. L., Tapley, I., Milne, A. K., Williams, M. L., Zhou, Z. S., Lehmann, E., Caccetta, P., Lowell, K., & Held, A. (2014). C- and L-band SAR interoperability: Filling the gaps in continuous forest cover mapping in Tasmania. *Remote Sensing of Environment*, *155*, 58–68. https://doi.org/10.1016/j.rse.2014.02.020

Momo Takoudjou, S., Ploton, P., Sonké, B., Hackenberg, J., Griffon, S., de Coligny, F., Kamdem, N. G., Libalah, M., Mofack, G., Le Moguédec, G., Pélissier, R., & Barbier, N. (2018). Using terrestrial laser scanning data to estimate large tropical trees biomass and calibrate allometric models: A comparison with traditional destructive approach. *Methods in Ecology and Evolution*, *9*(4), 905–916. https://doi.org/10.1111/2041-210X.12933

Müller, M. (2007). Dynamic Time Warping. In *Information retrieval for music and motion* (pp. 69–84). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-74048-3_4

Nielsen, A. A., Skriver, H., & Conradsen, K. (2007). Complex Wishart Distribution Based Analysis of Polarimetric Synthetic Aperture Radar Data. *2007 International Workshop on the Analysis of Multi-temporal Remote Sensing Images*, 1–6.

Nilsson, M. (1996). Estimation of tree heights and stand volume using an airborne lidar system. *Remote Sensing of Environment*, *56*(1), 1–7. https://doi.org/10.1016/0034-4257(95)00224-3

Norway's international climate and forest initiative (nicfi). (2021). https://www.nicfi.no/

Nuthammachot, N., Askar, A., Stratoulias, D., & Wicaksono, P. (2020). Combined use of Sentinel-1 and Sentinel-2 data for improving above-ground biomass estimation. *Geocarto International*, *0*(0), 1–11. https://doi.org/10.1080/10106049.2020.1726507

Olsen, N. L., Markussen, B., & Raket, L. L. (2018). Simultaneous inference for misaligned multivariate functional data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *67*(5), 1147–1176. https://doi.org/https://doi.org/10.1111/rssc.12276

Pekkarinen, A., Reithmaier, L., & Strobl, P. (2009). Pan-European forest/non-forest mapping with Landsat ETM+ and CORINE Land Cover 2000 data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *64*(2), 171–183. https://doi.org/10.1016/j.isprsjprs.2008.09.004

Pinty, B., & Verstraete, M. M. (1992). GEMI: a non-linear index to monitor global vegetation from satellites. *Vegetation*, *101*(1), 15–20. https://doi.org/10.1007/BF00031911

Pohl, C., & Van Genderen, J. L. (1998). Review article - Multisensor image fusion in remote sensing: concepts, methods and applications. *Earth*, *19*(5), 823–854.

Pourshamsi, M., Garcia, M., Lavalle, M., & Balzter, H. (2018). A Machine-Learning Approach to PolInSAR and LiDAR Data Fusion for Improved Tropical Forest Canopy Height Estimation Using NASA AfriSAR Campaign Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(10), 3453–3463. https://doi.org/10.1109/JSTARS.2018.2868119

Pourshamsi, M., Xia, J., Yokoya, N., Garcia, M., Lavalle, M., Pottier, E., & Balzter, H. (2021). Tropical forest canopy height estimation from combined polarimetric SAR and LiDAR using machine-learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, *172*(May 2020), 79–94. https://doi.org/10.1016/j.isprsjprs.2020.11.008

Qi, J., Chehbouni, A., Huete, A., Kerr, Y., & Sorooshian, S. (1994). A modified soil adjusted vegetation index. *Remote Sensing of Environment*, *48*(2), 119–126. https://doi.org/10.1016/0034-4257(94)90134-1

Qin, Y., Xiao, X., Wigneron, J. P., Ciais, P., Brandt, M., Fan, L., Li, X., Crowell, S., Wu, X., Doughty, R., Zhang, Y., Liu, F., Sitch, S., & Moore, B. (2021). Carbon loss from forest degradation exceeds that from deforestation in the Brazilian Amazon. *Nature Climate Change*, *11*(5), 442–448. https://doi.org/10.1038/s41558-021-01026-5

Quegan, S., Le Toan, T., Chave, J., Dall, J., Exbrayat, J. F., Minh, D. H. T., Lomas, M., D'Alessandro, M. M., Paillou, P., Papathanassiou, K., Rocca, F., Saatchi, S., Scipal, K., Shugart, H., Smallman, T. L., Soja, M. J., Tebaldini, S., Ulander, L., Villard, L., & Williams, M. (2019). The European Space Agency BIOMASS mission: Measuring forest above-ground biomass from space. *Remote Sensing of Environment*, *227*(April 2019), 44–60. https://doi.org/10.1016/j.rse.2019.03.032

Quegan, S., Toan, T. L., Yu, J. J., Ribbes, F., & Floury, N. (2000). Multitemporal ERS SAR analysis applied to forest mapping. *IEEE Transactions on Geoscience and Remote Sensing*, *38*(2 I), 741–753. https://doi.org/10.1109/36.842003

Quegan, S., & Yu, J. J. (2001). Filtering of multichannel SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(11), 2373–2379. https://doi.org/10.1109/36.964973

Rabiner, L., & Juang, B.-H. (1993). *Fundamentals of speech recognition*. PTR Prentice Hall.

Rahman, M. M., & Sumantyo, J. T. S. (2010). Mapping tropical forest cover and deforestation using synthetic aperture radar (SAR) images. *Applied Geomatics*, *2*(3), 113–121. https://doi.org/10.1007/s12518-010-0026-9

Raspaud, M., Hoese, D., Lahtinen, P., Finkensieper, S., Holl, G., Proud, S., Dybbroe, A., Meraner, A., Feltz, J., Zhang, X., Strandgren, Joro, S., Roberts, W., Rasmussen, L. Ø., BENR0, Méndez, J. H. B., Yufei Zhu, Mherbertson, Rdaruwala, . . . ColinDuff. (2022). Pytroll/satpy: Version 0.36.0 (2022/04/14). https://doi.org/10.5281/ZENODO.6460193

Rauste, Y., Antropov, O., Häme, T., Ramminger, G., Gomez, S., & Seifert, F. M. (2013). Mapping selective logging in tropical forest with space-borne SAR data. *ESA Special Publication*, *722*, 168.

Reiche, J., de Bruin, S., Hoekman, D., Verbesselt, J., & Herold, M. (2015). A Bayesian Approach to Combine Landsat and ALOS PALSAR Time Series for Near Real-Time Deforestation Detection. *Remote Sensing*, *7*(5), 4973–4996. https://doi.org/10.3390/rs70504973

Reiche, J., Hamunyela, E., Verbesselt, J., Hoekman, D., & Herold, M. (2018). Improving near-real time deforestation monitoring in tropical dry forests by combining dense Sentinel-1 time series with Landsat and ALOS-2 PALSAR-2. *Remote Sensing of Environment*, *204*(November 2017), 147–161. https://doi.org/10.1016/j.rse.2017.10.034

Reiche, J., Lucas, R., Mitchell, A. L., Verbesselt, J., Hoekman, D. H., Haarpaintner, J., Kellndorfer, J. M., Rosenqvist, A., Lehmann, E. A., Woodcock, C. E., Seifert, F. M., & Herold, M. (2016). Combining satellite data for better tropical forest monitoring. *Nature Climate Change*, *6*(2), 120–122. https://doi.org/10.1038/nclimate2919

Reiche, J., Souza, C. M., Hoekman, D. H., Verbesselt, J., Persaud, H., & Herold, M. (2013). Feature Level Fusion of Multi-Temporal ALOS PALSAR and Landsat Data for Mapping and Monitoring of Tropical Deforestation and Forest Degradation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *6*(5), 2159–2173. https://doi.org/10.1109/JSTARS.2013.2245101

Reiche, J., Verbesselt, J., Hoekman, D., & Herold, M. (2015). Fusing Landsat and SAR time series to detect deforestation in the tropics. *Remote Sensing of Environment*, *156*, 276–293. https://doi.org/10.1016/j.rse.2014.10.001

Riofrio, S., Pozo, D., Rosero, J., & Vasquez, J. (2017). Gesture Recognition Using Dynamic Time Warping and Kinect: A Practical Approach, 302–308. https://doi.org/10.1109/INCISCOS.2017.36

Roberts, J., Balzter, H., Yaqing Gou, Louis, V., & Robb, C. (2020). Pyeo: Automated satellite imagery processing. https://doi.org/10.5281/ZENODO.3689674

Rodríguez-Veiga, P., Wheeler, J., Louis, V., Tansey, K., & Balzter, H. (2017). Quantifying Forest Biomass Carbon Stocks From Space. *Current Forestry Reports*, *3*(1), 1–18. https://doi.org/10.1007/s40725-017-0052-5

Rondeaux, G., Steven, M., & Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment*, *55*(2), 95–107. https://doi.org/10.1016/0034-4257(95)00186-7

Rosich, B., & Meadows, P. (2004). *Absolute calibration of ASAR Level 1 products generated with PF-ASAR* (Technical Note No. 1) [Reference ENVI-CLVL-EOPG-TN-03-0010]. European Space Agency. https://earth.esa.int/web/guest/-/absolute-calibration-of-asar-level-1-products-generated-with-pf-asar-4503

Rouse, W., Haas, H., & Deering, W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *351*, 309. https://ntrs.nasa.gov/search.jsp?R=19740022614

Sakoe, H., & Chiba, S. (1978). Dynamic Programming Algorithm Optimization for Spoken Word Recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *26*(1), 43–49. https://doi.org/10.1109/TASSP.1978.1163055

Salvador, S., & Chan, P. (2007). Toward accurate dynamic time warping in linear time and space. *Intelligent Data Analysis*, *11*(5), 561–580.

Santoro, M., Cartus, O., Carvalhais, N., Rozendaal, D. M. A., Avitabile, V., Araza, A., de Bruin, S., Herold, M., Quegan, S., Rodr\'\iguez-Veiga, P., Balzter, H., Carreiras, J., Schepaschenko, D., Korets, M., Shimada, M., Itoh, T., Moreno Mart\'\inez, Á., Cavlovic, J., Cazzolla Gatti, R., . . . Willcock, S. (2021). The global forest above-ground biomass pool for 2010 estimated from high-resolution satellite observations. *Earth System Science Data*, *13*(8), 3927–3950. https://doi.org/10.5194/essd-13-3927-2021

Scheiber, R., & Moreira, A. (2000). Coregistration of interferometric SAR images using spectral diversity. *IEEE Transactions on Geoscience and Remote Sensing*, *38*(5 I), 2179–2191. https://doi.org/10.1109/36.868876

Schmidt, G., Jenkerson, C., Masek, J., Vermote, E., & Gao, F. (2013). *Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) Algorithm Description* (tech. rep. December). US Geological Survey. http://pubs.usgs.gov/of/2013/1057/

Shimada, M., Isoguchi, O., Tadono, T., & Isono, K. (2009). PALSAR radiometric and geometric calibration. *IEEE Transactions on Geoscience and Remote Sensing*, *47*(12), 3915–3932. https://doi.org/10.1109/TGRS.2009.2023909

Shimada, M., Itoh, T., Motooka, T., Watanabe, M., Shiraishi, T., Thapa, R., & Lucas, R. (2014). New global forest/non-forest maps from ALOS PALSAR data (2007-2010). *Remote Sensing of Environment*, *155*, 13–31. https://doi.org/10.1016/j.rse.2014.04.014

Shimada, M., & Ohtaki, T. (2010). Generating Large-Scale High-Quality SAR Mosaic Datasets: Application to PALSAR Data for Global Monitoring. *IEEE*

*Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *3*(4), 637–656. https://doi.org/10.1109/JSTARS.2010.2077619

Skutsch, M., Torres, A. B., Mwampamba, T. H., Ghilardi, A., & Herold, M. (2011). Dealing with locally-driven degradation: A quick start option under REDD+. *Carbon Balance and Management*, *6*, 1–7. https://doi.org/10.1186/1750-0680-6-16

Small, D. (2011). Flattening gamma: Radiometric terrain correction for SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *49*(8), 3081–3093. https://doi.org/10.1109/TGRS.2011.2120616

Small, D., & Schubert, A. (2008). Guide to ASAR Geocoding. (1.01), 36.

Soja, M. J., Persson, H. J., & Ulander, L. M. (2018). Modeling and Detection of Deforestation and Forest Growth in Multitemporal TanDEM-X Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(10), 3548–3563. https://doi.org/10.1109/JSTARS.2018.2851030

Soja, M. J., Quegan, S., D'Alessandro, M. M., Banda, F., Scipal, K., Tebaldini, S., & Ulander, L. M. (2021). Mapping above-ground biomass in tropical forests with ground-cancelled P-band SAR and limited reference data. *Remote Sensing of Environment*, *253*, 112153. https://doi.org/10.1016/j.rse.2020.112153

Solberg, S., Astrup, R., Breidenbach, J., Nilsen, B., & Weydahl, D. (2013). Monitoring spruce volume and biomass with InSAR data from TanDEM-X. *Remote Sensing of Environment*, *139*, 60–67. https://doi.org/10.1016/j.rse.2013.07.036

Solberg, S., May, J., Bogren, W., Breidenbach, J., Torp, T., & Gizachew NIBIO, B. (2015). Forest carbon change in Uganda 2000-2012 estimated with InSAR. *2012*(October). http://lps16.esa.int/posterfiles/paper1369/Forest%20carbon%20change%20in%20Uganda%202000-2012.pdf

Souza, C. M., Roberts, D. A., & Cochrane, M. A. (2005). Combining spectral and spatial information to map canopy damage from selective logging and forest fires. *Remote Sensing of Environment*, *98*(2-3), 329–343. https://doi.org/10.1016/j.rse.2005.07.013

Souza, C. M., & Siqueira, J. (2013). ImgTools: a software for optical remotely sensed data analysis. *Anais XVI Simpósio Brasileiro de Sensoriamento Remoto*, 1571–1578. http://www.dsr.inpe.br/sbsr2013/files/p1235.pdf

Sripada, R. P., Heiniger, R. W., White, J. G., & Meijer, A. D. (2006). Aerial color infrared photography for determining early in-season nitrogen requirements in corn. *Agronomy Journal*, *98*(4), 968–977. https://doi.org/10.2134/agronj2005.0200

Strozzi, T., Dammert, P., Wegmüller, U., Martinez, J. M., Beaudoin, A., Askne, J., & Hallikainen, M. (1998). European forest mapping with SAR interferometry. *Earth Observation Quarterly*, (62), 17–20.

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, *8*(2), 127–150. https://doi.org/10.1016/0034-4257(79)90013-0

Vancutsem, C., Achard, F., Pekel, J. F., Vieilledent, G., Carboni, S., Simonetti, D., Gallego, J., Aragão, L. E., & Nasi, R. (2021). Long-term (1990–2019) monitoring of forest cover changes in the humid tropics. *Science Advances*, *7*(10), 1–22. https://doi.org/10.1126/sciadv.abe1603

Verbesselt, J., Hyndman, R., Newnham, G., & Culvenor, D. (2010). Detecting trend and seasonal changes in satellite image time series. *Remote Sensing of Environment*, *114*(1), 106–115. https://doi.org/10.1016/j.rse.2009.08.014

Verbesselt, J., Hyndman, R., Zeileis, A., & Culvenor, D. (2010). Phenological change detection while accounting for abrupt and gradual trends in satellite image time series. *Remote Sensing of Environment*, *114*(12), 2970–2980. https://doi.org/10.1016/j.rse.2010.08.003

Verbesselt, J., Zeileis, A., & Herold, M. (2012). Near real-time disturbance detection using satellite image time series. *Remote Sensing of Environment*, *123*, 98–108. https://doi.org/10.1016/j.rse.2012.02.022

Verhelst, K., Gou, Y., Herold, M., & Reiche, J. (2021). Improving forest baseline maps in tropical wetlands using gedi-based forest height information and sentinel-1. *Forests*, *12*(10), 1–16. https://doi.org/10.3390/f12101374

Vicente-Serrano, S. M., Gouveia, C., Camarero, J. J., Begueria, S., Trigo, R., Lopez-Moreno, J. I., Azorin-Molina, C., Pasho, E., Lorenzo-Lacruz, J., Revuelto, J., Moran-Tejeda, E., & Sanchez-Lorenzo, A. (2013). Response of vegetation to drought time-scales across global land biomes. *Proceedings of the National Academy of Sciences*, *110*(1), 52–57. https://doi.org/10.1073/pnas.1207068110

Wegmüller, U., & Werner, C. (1997). Retrieval of vegetation parameters with sar interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, *35*(1), 18–24.

Wegmüller, U., & Werner, C. L. (1995). Sar interferometric signatures of forest. *IEEE Transactions on Geoscience and Remote Sensing*, *33*(5), 1153–1161.

Werner, C., Wegmüller, U., Strozzi, T., & Wiesmann, A. (2000). GAMMA SAR and interferometric processing software. *Proceedings of the ERS-Envisat Symposium, Gothenburg, Sweden*, (461), 211–219.

Woodhouse, I. H. (2005). *Introduction to Microwave Remote Sensing*.

Woodhouse, I. H., Mitchard, E. T. A., Brolly, M., Maniatis, D., & Ryan, C. M. (2012). Radar backscatter is not a 'direct measure' of forest biomass. *Nature Climate Change*, *2*(8), 556–557. https://doi.org/10.1038/nclimate1601

Zhang, X., Friedl, M. A., & Schaaf, C. B. (2009). Sensitivity of vegetation phenology detection to the temporal resolution of satellite data. *International Journal of Remote Sensing*, *30*(8), 2061–2074. https://doi.org/10.1080/01431160802549237

Zhu, Z., & Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, *118*, 83–94. https://doi.org/10.1016/j.rse.2011.10.028

Zhu, Z., & Woodcock, C. E. (2014a). Automated cloud, cloud shadow, and snow detection in multitemporal Landsat data: An algorithm designed specifically for monitoring land cover change. *Remote Sensing of Environment*, *152*, 217–234. https://doi.org/10.1016/j.rse.2014.06.012

Zhu, Z., & Woodcock, C. E. (2014b). Continuous change detection and classification of land cover using all available Landsat data. *Remote Sensing of Environment*, *144*, 152–171. https://doi.org/10.1016/j.rse.2014.01.011

Zhu, Z., Woodcock, C. E., & Olofsson, P. (2012). Continuous monitoring of forest disturbance using all available Landsat imagery. *Remote Sensing of Environment*, *122*, 75–91. https://doi.org/10.1016/j.rse.2011.10.030

Zimbres, B., Rodríguez-Veiga, P., Shimbo, J. Z., da Conceição Bispo, P., Balzter, H., Bustamante, M., Roitman, I., Haidar, R., Miranda, S., Gomes, L., Alvim Carvalho, F., Lenza, E., Maracahipes-Santos, L., Abadia, A. C., do Prado Júnior, J. A., Mendonça Machado, E. L., Dias Gonzaga, A. P., de Castro Nunes Santos Terra, M., de Mello, J. M., … Alencar, A. (2021). Mapping the stock and spatial distribution of aboveground woody biomass in the native vegetation of the Brazilian Cerrado biome. *Forest Ecology and Management*, *499*, 119615. https://doi.org/https://doi.org/10.1016/j.foreco.2021.119615

# Further Reading

Abbas, S., Wong, M. S., Wu, J., Shahzad, N., & Irteza, S. M. (2020). Approaches of satellite remote sensing for the assessment of above-ground biomass across tropical forests: Pan-tropical to national scales. *Remote Sensing*, *12*(20), 1–38. https://doi.org/10.3390/rs12203351

Adams, R. P., & MacKay, D. J. C. (2007). Bayesian Online Changepoint Detection. https://doi.org/arXiv:0710.3742v1

Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, *18*(1), 1–22. https://doi.org/10.1002/jae.659

Bazi, Y. (2004). Change detection in multitemporal SAR images based on generalized Gaussian distribution and EM algorithm. *Proceedings of SPIE*, *5573*, 364–375. https://doi.org/10.1117/12.567890

Bazi, Y., Bruzzone, L., & Melgani, F. (2005). An unsupervised approach based on the generalized Gaussian model to automatic change detection in multitemporal SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, *43*(4), 874–886. https://doi.org/10.1109/TGRS.2004.842441

Birth, G. S., & McVey, G. R. (1968). Measuring the Color of Growing Turf with a Reflectance Spectrophotometer 1. *Agronomy Journal*, *60*(6), 640–643.

Bush, A., Sollmann, R., Wilting, A., Bohmann, K., Cole, B., Balzter, H., Martius, C., Zlinszky, A., Calvignac-Spencer, S., Cobbold, C. A., Dawson, T. P., Emerson, B. C., Ferrier, S., Gilbert, M. T. P., Herold, M., Jones, L., Leendertz, F. H., Matthews, L., Millington, J. D. A., . . . Yu, D. W. (2017). Connecting Earth observation to high-throughput biodiversity data. *Nature Ecology & Evolution*, *1*(7), 176. https://doi.org/10.1038/s41559-017-0176

Carincotte, C., Derrode, S., & Bourennane, S. (2006). Unsupervised change detection on SAR images using fuzzy hidden Markov chains. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(2), 432–441. https://doi.org/10.1109/TGRS.2005.861007

Cartus, O., Santoro, M., Wegmüller, U., & Rommen, B. (2019). Benchmarking the retrieval of biomass in boreal forests using P-band SAR backscatter with multi-temporal C- and L-band observations. *Remote Sensing*, *11*(14). https://doi.org/10.3390/rs11141695

Chang, Z., Hobeichi, S., Wang, Y. P., Tang, X., Abramowitz, G., Chen, Y., Cao, N., Yu, M., Huang, H., Zhou, G., Wang, G., Ma, K., Du, S., Li, S., Han, S., Ma, Y., Wigneron, J. P., Fan, L., Saatchi, S. S., & Yan, J. (2021). New forest aboveground biomass maps of China integrating multiple datasets. *Remote Sensing*, *13*(15), 1–20. https://doi.org/10.3390/rs13152892

Chatfield, C. (2016). *The Analysis of Time Series: An Introduction*.

Collins, M. B., & Mitchard, E. T. (2015). Integrated radar and lidar analysis reveals extensive loss of remaining intact forest on Sumatra 2007-2010. *Biogeosciences*, *12*(22), 6637–6653. https://doi.org/10.5194/bg-12-6637-2015

Davison, A. C. (2003). *Statistical Models*.

DeGroot, M. H., Schervish, M. J., & Sheet, C. (2011). *Probability and Statistics*. https://doi.org/0321709705

Dillon, R. J. (2013). Processing Remote Sensing Data with Python Sciences Processing Remote Sensing Data with Python.

Dostálová, A., Lang, M., Ivanovs, J., Waser, L. T., & Wagner, W. (2021). European wide forest classification based on sentinel-1 data. *Remote Sensing*, *13*(3), 1–27. https://doi.org/10.3390/rs13030337

Euillades, P., Euillades, L., Pepe, A., Mastro, P., Falabella, F., Imperatore, P., Tang, Y., & Rosell, P. (2021). Recent advancements in multi-temporal methods applied to new generation SAR systems and applications in South America. *Journal of South American Earth Sciences*, *111*(May), 103410. https://doi.org/10.1016/j.jsames.2021.103410

Fatoyinbo, T., Armston, J., Simard, M., Saatchi, S., Denbina, M., Lavalle, M., Hofton, M., Tang, H., Marselis, S., Pinto, N., Hancock, S., Hawkins, B., Duncanson, L., Blair, B., Hansen, C., Lou, Y., Dubayah, R., Hensley, S., Silva, C., . . . Hibbard, K. (2021). The NASA AfriSAR campaign: Airborne SAR and lidar measurements of tropical forest structure and biomass in support of current and future space missions. *Remote Sensing of Environment*, *264*(June 2020). https://doi.org/10.1016/j.rse.2021.112533

Fearnhead, P. (2006). Exact and efficient Bayesian inference for multiple change-point problems. *Statistics and computing*, *16*(2), 203–213.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995). *Bayesian data analysis*. Chapman; Hall/CRC.

Ghanbari, M., & Akbari, V. (2015). Generalized minimum-error thresholding for unsupervised change detection from multilook polarimetric SAR data. *International Geoscience and Remote Sensing Symposium (IGARSS)*, *2015-Novem*(10), 1853–1856. https://doi.org/10.1109/IGARSS.2015.7326153

Hamad, R., Balzter, H., & Kolo, K. (2018). Predicting Land Use/Land Cover Changes Using a CA-Markov Model under Two Different Scenarios. *Sustainability*, *10*(10). https://doi.org/10.3390/su10103421

Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., & Hobart, G. W. (2015). Regional detection, characterization, and attribution of annual forest change from 1984 to 2012 using Landsat-derived time-series metrics. *Remote Sensing of Environment*, *170*, 121–132. https://doi.org/10.1016/j.rse.2015.09.004

Hoekman, D., Kooij, B., Quiñones, M., Vellekoop, S., Carolita, I., Budhiman, S., Arief, R., & Roswintiarti, O. (2020). Wide-area near-real-time monitoring of tropical forest degradation and deforestation using sentinel-1. *Remote Sensing*, *12*(19), 1–32. https://doi.org/10.3390/rs12193263

Hoekman, D. H., & Quinones, M. J. (2000). Land cover type and biomass classification using AirSAR data for evaluation of monitoring scenarios in the Colombian Amazon. *IEEE Transactions on Geoscience and Remote Sensing*, *38*(2 I), 685–696. https://doi.org/10.1109/36.841998

Houghton, J. (1989). *Introduction to the physics and techniques of remote sensing* (Vol. 54). https://doi.org/10.1016/0031-9201(89)90201-X

Huete, A., Didan, K., Miura, T., Rodriguez, E., Gao, X., & Ferreira, L. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, *83*(1-2), 195–213. https://doi.org/10.1016/S0034-4257(02)00096-2

Ibrahim, S., Kaduk, J., Tansey, K., Balzter, H., & Lawal, U. M. (2021). Detecting phenological changes in plant functional types over West African savannah dominated landscape. *International Journal of Remote Sensing*, *42*(2), 567–594. https://doi.org/10.1080/01431161.2020.1811914

Lehmann, E. A., Wallace, J. F., Caccetta, P. A., Furby, S. L., & Zdunic, K. (2012). Forest cover trends from time series Landsat data for the Australian continent. *International Journal of Applied Earth Observation and Geoinformation*, *21*(1), 453–462. https://doi.org/10.1016/j.jag.2012.06.005

Leisch, F., Hornik, K., & Kuan, C. M. (2000). Monitoring structural changes with the generalized fluctuation test. *Econometric Theory*, *16*(6), 835–854. https://doi.org/10.1017/S0266466600166022

Lenth, R. V., & Lenth, R. V. (2016). Some Practical Guidelines for Effective Sample Size Determination Published by : Taylor & Francis , Ltd . on behalf of the American Statistical Association Stable URL : http://www.jstor.org/stable/2685797 Your use of the JSTOR archive indicates your accep. *55*(3), 187–193.

Lhermitte, S., Verbesselt, J., Verstraeten, W. W., & Coppin, P. (2011). A comparison of time series similarity measures for classification and change detection of ecosystem dynamics. *Remote Sensing of Environment*, *115*(12), 3129–3152. https://doi.org/10.1016/j.rse.2011.06.020

Lillesand, T. M., Kiefer, R. W., & Chipman, J. W. (1989). *Remote Sensing and Image Interpretation* (Vol. 53). https://doi.org/10.1017/CBO9781107415324.004

Longépé, N., Rakwatin, P., Isoguchi, O., Shimada, M., Uryu, Y., & Yulianto, K. (2011). Assessment of ALOS PALSAR 50 m orthorectified FBD data for regional land cover classification by support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, *49*(6 PART 1), 2135–2150. https://doi.org/10.1109/TGRS.2010.2102041

Marino, A., Cloude, S. R., & Lopez-Sanchez, J. M. (2013). A new polarimetric change detector in radar imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *51*(5), 2986–3000. https://doi.org/10.1109/TGRS.2012.2211883

Menlove, J., & Healey, S. P. (2020). A comprehensive forest biomass dataset for the USA allows customized validation of remotely sensed biomass estimates. *Remote Sensing*, *12*(24), 1–13. https://doi.org/10.3390/rs12244141

Miettinen, J., Stibig, H. J., & Achard, F. (2014). Remote sensing of forest degradation in Southeast Asia-Aiming for a regional view through 5-30 m satellite data. *Global Ecology and Conservation*, *2*, 24–36. https://doi.org/10.1016/j.gecco.2014.07.007

Morton, D. C., DeFries, R. S., Nagol, J., Souza, C. M., Kasischke, E. S., Hurtt, G. C., & Dubayah, R. (2011). Mapping canopy damage from understory fires in Amazon forests using annual time series of Landsat and MODIS data. *Remote Sensing of Environment*, *115*(7), 1706–1720. https://doi.org/10.1016/j.rse.2011.03.002

Motohka, T., Shimada, M., Uryu, Y., & Setiabudi, B. (2014). Using time series PALSAR gamma nought mosaics for automatic detection of tropical deforestation: A test study in Riau, Indonesia. *Remote Sensing of Environment*, *155*, 79–88. https://doi.org/10.1016/j.rse.2014.04.012

Murphy, K. (2007). Modeling Changing Dependency Structure in Multivariate Time Series. *Proceedings of the 24th Annual International Conference on Machine Learning (ICML 2007)*, 1055–1062.

Murphy, K. R., Myors, B., & Wolach, A. (2014). *Statistical power analysis: A simple and general model for traditional and modern hypothesis tests*.

Peter, H., Jäggi, A., Fernández, J., Escobar, D., Ayuga, F., Arnold, D., Wermuth, M., Hackel, S., Otten, M., Simons, W., Visser, P., Hugentobler, U., & Féménias, P. (2017). Sentinel-1A – First precise orbit determination results. *Advances in Space Research*, *60*(5), 879–892. https://doi.org/10.1016/J.ASR.2017.05.034

Plaza, A. J., & Chang, C. (2007). *High Performance Computing in Remote Sensing*.

Potapov, P. V., Turubanova, S. A., Hansen, M. C., Adusei, B., Broich, M., Altstatt, A., Mane, L., & Justice, C. O. (2012). Quantifying forest cover loss in Democratic Republic of the Congo, 2000-2010, with Landsat ETM+ data. *Remote Sensing of Environment*, *122*, 106–116. https://doi.org/10.1016/j.rse.2011.08.027

Qi, W., & Dubayah, R. O. (2016). Combining Tandem-X InSAR and simulated GEDI lidar observations for forest structure mapping. *Remote Sensing of Environment*, *187*(2016), 253–266. https://doi.org/http://dx.doi.org/10.1016/j.rse.2016.10.018

Rakwatin, P., Longépé, N., Isoguchi, O., Shimada, M., Uryu, Y., & Takeuchi, W. (2012). Using multiscale texture information from ALOS PALSAR to map tropical forest. *International Journal of Remote Sensing*, *33*(24), 7727–7746. https://doi.org/10.1080/01431161.2012.701349

Rangayyan, R. M., Acha, B., Serrano, C., Joshi, P., Marco, F., van der Spoel, E., Rozing, M. P., Houwing-Duistermaat, J. J., Eline Slagboom, P., Beekman, M., de Craen, A. J. M., Westendorp, R. G. J., van Heemst, D., A. Zisserman, & Jenkins, W. K. (2016). Fourier Series, Fourier Transforms, and the DFT. *Information Engineering, Oxford University*, *7*(11), 956–963. https://doi.org/10.1117/3.887920

Rees, W. G. (2013). *Physical principles of remote sensing*. Cambridge University Press.

Richards, J., & Jia, X. (2006). *Remote Sensing Digital Image Analysis*. https://doi.org/10.1016/S0169-555X(01)00164-7

Rignot, E., & van Zyl, J. (1993). Change detection techniques for ERS-1 SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, *31*(4), 896–906. https://doi.org/10.1109/36.239913

Rodríguez-Veiga, P., Saatchi, S., Tansey, K., & Balzter, H. (2016). Magnitude, spatial distribution and uncertainty of forest biomass stocks in Mexico. *Remote Sensing of Environment*, *183*, 265–281. https://doi.org/10.1016/j.rse.2016.06.004

Sarker, M. L. R., Nichol, J., Iz, H. B., Ahmad, B. B., & Rahman, A. A. (2013). Forest biomass estimation using texture measurements of high-resolution dual-polarization C-band SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, *51*(6), 3371–3384. https://doi.org/10.1109/TGRS.2012.2219872

Schroeder, T. A., Healey, S. P., Moisen, G. G., Frescino, T. S., Cohen, W. B., Huang, C., Kennedy, R. E., & Yang, Z. (2014). Improving estimates of forest disturbance by combining observations from landsat time series with U.S. Forest Service Forest Inventory and Analysis data. *Remote Sensing of Environment*, *154*(1), 61–73. https://doi.org/10.1016/j.rse.2014.08.005

Shan, Z., Wang, C., Zhang, H., & Chen, J. (2011). H-$\alpha$ Decomposition and Alternative Parameters for Dual Polarization SAR Data. *PIERS Proc.*, 1386–1390.

Shiraishi, T., Motohka, T., Thapa, R. B., Watanabe, M., & Shimada, M. (2014). Comparative assessment of supervised classifiers for land use-land cover classification in a tropical region using time-series PALSAR mosaic data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *7*(4), 1186–1199. https://doi.org/10.1109/JSTARS.2014.2313572

Shokoohi-Yekta, M., Wang, J., & Keogh, E. (2015). On the non-trivial generalization of Dynamic Time Warping to the multi-dimensional case. *SIAM International Conference on Data Mining 2015, SDM 2015*, 289–297. https://doi.org/10.1137/1.9781611974010.33

Silva, C. A., Duncanson, L., Hancock, S., Neuenschwander, A., Thomas, N., Hofton, M., Fatoyinbo, L., Simard, M., Marshak, C. Z., Armston, J., Lutchke, S., & Dubayah, R. (2021). Fusing simulated GEDI, ICESat-2 and NISAR data for regional aboveground biomass mapping. *Remote Sensing of Environment*, *253*(May 2020), 112234. https://doi.org/10.1016/j.rse.2020.112234

Souza, C., Firestone, L., Silva, L. M., & Roberts, D. (2003). Mapping forest degradation in the Eastern Amazon from SPOT 4 through spectral mixture models. *Remote Sensing of Environment*, *87*(4), 494–506. https://doi.org/10.1016/j.rse.2002.08.002

Thapa, R. B., Itoh, T., Shimada, M., Watanabe, M., Takeshi, M., & Shiraishi, T. (2014). Evaluation of ALOS PALSAR sensitivity for characterizing natural forest cover in wider tropical areas. *Remote Sensing of Environment*, *155*, 32–41. https://doi.org/10.1016/j.rse.2013.04.025

Thompson, S. K. (1945). *Sampling*. https://doi.org/http://dx.doi.org/10.1002/9781118162934

Trouve, E., Chambenoit, Y., Classeau, N., & Bolon, P. (2003). Statistical and Operational Performance Assessment of Multitemporal SAR Image Filtering. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(11 PART I), 2519–2530. https://doi.org/10.1109/TGRS.2003.817270

Verbesselt, J., Umlauf, N., Hirota, M., Holmgren, M., Van Nes, E. H., Herold, M., Zeileis, A., & Scheffer, M. (2016). Remotely sensed resilience of tropical forests. *Nature Climate Change*, *6*(11), 1028–1031. https://doi.org/10.1038/nclimate3108

Walter Zucchini, L. L. M. (2009). Hidden Markov Models for Time series, 64–66.

Wang, J. F., Stein, A., Gao, B. B., & Ge, Y. (2012). A review of spatial sampling. *Spatial Statistics*, *2*(1), 1–14. https://doi.org/10.1016/j.spasta.2012.08.001

Watts, L. M., & Laffan, S. W. (2014). Effectiveness of the BFAST algorithm for detecting vegetation response patterns in a semi-arid region. *Remote Sensing of Environment*, *154*(1), 234–245. https://doi.org/10.1016/j.rse.2014.08.023

Wiederkehr, N. C., Gama, F. F., Castro, P. B., Bispo, P. d. C., Balzter, H., Sano, E. E., Liesenberg, V., Santos, J. R., & Mura, J. C. (2020). Discriminating forest successional stages, forest degradation, and land use in central Amazon using ALOS/PALSAR-2 full-polarimetric data. *Remote Sensing*, *12*(21), 1–30. https://doi.org/10.3390/rs12213512

Williams, M., Milne, A., & Mitchell, A. (2015). Towards operational forest monitoring using satellite Synthetic Aperture Radar. *International Conference on Space Science and Communication, IconSpace*, *2015-Septe*, 209–214. https://doi.org/10.1109/IconSpace.2015.7283790

Zeileis, A. (2005). A unified approach to structural change tests based on ML scores, F statistics, and OLS residuals. *Econometric Reviews*, *24*(4), 445–466. https://doi.org/10.1080/07474930500406053

Zeileis, A., & Kleiber, C. (2005). Validating multiple structural change models-a case study. *Journal of Applied Econometrics*, *20*(5), 685–690. https://doi.org/10.1002/jae.856

Zenz, G., & Goldgruber, M. (2013). Benchmark for success. *International Water Power and Dam Construction*, *65*(8), 30–32. https://doi.org/10.1073/pnas.1019576108

Zhang, Y., & Zhao, H. (2020). Land–Use and Land-Cover Change Detection Using Dynamic Time Warping–Based Time Series Clustering Method. *Canadian Journal of Remote Sensing*, *46*(1), 67–83. https://doi.org/10.1080/07038992.2020.1740083

Zhu, Z. (2017). Change detection using landsat time series: A review of frequencies, preprocessing, algorithms, and applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, *130*, 370–384. https://doi.org/10.1016/j.isprsjprs.2017.06.013

Zimbres, B., Rodriguez-Veiga, P., Shimbo, J. Z., da Conceicao Bispo, P., Balzter, H., Bustamante, M., Roitman, I., Haidar, R., Miranda, S., Gomes, L., Alvim, F., Lenza, E., Maracahipes-Santos, L., Abadia, A. C., Junior, J. P., Machado, E. L. M., Gonzaga, A. P. D., de Castro Nunes Santos Terra, M., De Mello, J. M., . . . Alencar, A. (2021). Aboveground Woody Biomass Estimation of the Brazilian Cerrado Biome Using Data Integration, 1402–1405. https://doi.org/10.1109/igarss47720.2021.9554389