Investigation and application of temporal modelling techniques for large-scale mangrove monitoring

Katie Awty-Carroll

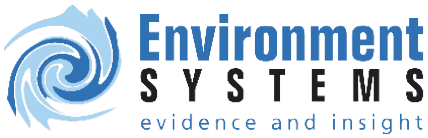
The Department of Geography and Earth Sciences

Aberystwyth University

A thesis submitted for the degree of

Doctor of Philosophy

December 2021



Knowledge Economy Skills Scholarships (KESS 2) is a pan-Wales higher level skills initiative led by Bangor University on behalf of the HE sector in Wales. It is part funded by the Welsh Government's European Social Fund (ESF) convergence programme for West Wales and the Valleys.

Statement of Originality

Declaration

Word count of thesis: 61,029

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed ... . (candidate)

Date .....20th May 2022.....

Statement 1

This thesis is the result of my own investigations, except where otherwise stated. Where correction services have been used, the extent and nature of the correction is clearly marked in a footnote(s).

Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed ... . (candidate)

Date .....20th May 2022.....

Statement 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ... . (candidate)

Date.....20th May 2022.....

Acknowledgements

Firstly I would like to thank my supervisors, Pete Bunting and Andy Hardy, for all of their support, advice, and enthusiasm throughout the project. I would also like to thank Gemma Bell and Environment Systems, for part-funding the project and offering a much needed different perspective.

My thanks also go to my fellow PhD students, Lilliana, Elikana, Maddie, Guy, George, and Osian, for making the office such a great place to work and for sharing their Python expertise.

The analyses in this thesis were made possible by the Supercomputing Wales project, and particular thanks goes to Colin Sauze, for his advice on scaling up my processing and for not minding too much when I almost broke Sunbird. I would also like to extend my thanks to all of the individuals in the open source community who spent their time writing the libraries of code upon which this thesis is built - GDAL, scikit-learn, the Open Data Cube, and many, many others.

Without funding from the European Social Fund and KESS 2, this project would never have existed. I would like to thank the KESS team for all their help in not making the admin too arduous, and for always being happy to answer my questions.

I would also like to thank my colleagues at Plymouth Marine Laboratory for their support and encouragement over the last few months.

Finally, special thanks go to my partner Danny, for never getting bored of talking about mangroves. And Matilda, of course – for being, almost literally, a rock.

Abstract

Recent advances in satellite observation have allowed for increasingly detailed monitoring of the Earth's surface, providing the opportunity to capture the complex seasonal dynamics of many Land Cover (LC) types and to detect LC change more rapidly than ever before. In addition, capacity for large-scale LC monitoring applications is constantly increasing due to advances in computing capability. As a result there has been an increase in methods that monitor LC change on a per-acquisition rather than a yearly basis, such as Continuous Change Detection and Classification (CCDC). This thesis utilises the Landsat data record to investigate and apply a time series modelling method to the problem of mangrove forest change. Mangroves are of high biological, economic, and ecological importance globally and are particularly vulnerable to the effects of climate change in addition to being threatened by over-exploitation, pollution, and expansion of aquaculture practices. Long-term monitoring of global mangrove populations is therefore vital, and impossible to achieve without repeated satellite imagery. However, data availability from satellites for tropical and sub-tropical areas is often limited due to cloud cover, and the dynamic nature of mangrove ecosystems introduces uncertainty for traditional monitoring approaches. These problems can be mitigated using a season-trend modelling approach to utilise every available observation on a pixel-wise basis, accounting for ephemeral change and removing the need for whole cloud free images.

A methodology for generating simulated optical time series is developed and used to objectively assess the ability of four algorithms to detect both the timing and nature of different LC change types. Based on this assessment, CCDC was chosen as being the most suitable method for monitoring of mangrove ecosystems due to its robustness to missing data and low commission error. The CCDC algorithm was then implemented and applied within a High Performance Computing framework. Using CCDC, yearly class maps were generated for six study sites for the last 30 years, with the method achieving an overall classification accuracy >90% and providing the most comprehensive assessment yet of mangrove extent in the Sundarbans and Niger Delta regions. Results showed that while mangrove extent in the Sundarbans has remained stable, nearly a quarter of the forest shows evidence of degradation. In addition, a trend analysis found that 11% of the Sundarbans was affected by the impact of Cyclone Sidr in 2007, 48% of which had not recovered by mid-2018. For the Niger Delta, the method achieved high accuracy from the 2000s onwards despite extremely low data availability. Observation of extent over time suggests that CCDC was also able to capture changes in extent caused by the 2015 mangrove die-back event in the Gulf of Carpentaria, Northern Australia and highlighted a net loss of mangroves in the Matang Forest Reserve over the last two decades, despite ongoing management. CCDC is therefore a promising methodology for global, long-term monitoring of mangroves, allowing for broader changes in extent to be examined in addition to providing details of mangrove condition change. The thesis concludes with a discussion of each chapter within the broader context of Earth Observation science, and identifies areas for future research.

List of publications

Chapter 3

Awty-Carroll, K.; Bunting, P.; Hardy, A.; Bell, G. An Evaluation and Comparison of Four Dense Time Series Change Detection Methods Using Simulated Data. *Remote Sensing.* **2019**, 11, 2779.

Chapter 4

Awty-Carroll, K.; Bunting, P.; Hardy, A.; Bell, G. Using Continuous Change Detection and Classification of Landsat Data to Investigate Long-Term Mangrove Dynamics in the Sundarbans Region. *Remote Sensing.* **2019**, 11, 2833.

Chapter 5

Awty-Carroll, K.; Bunting, P.; Hardy, A.; Bell, G. Evaluation of the Continuous Monitoring of Land Disturbance algorithm for large-scale mangrove classification. *Remote Sensing.* **2021**, 13, 3978.

Contents

List of Figures

List of Tables

List of Abbreviations

ALOS Advanced Land Observation Satellite

ANOVA Analysis of Variance

ARIMA Autoregressive Integrated Moving Average

AVHRR Advanced Very High Resolution Radiometer

AWS Amazon Web Services

BAP Best Available Pixel

BFAST Breaks for Additive and Seasonal Trend

BIC Bayesian Information Criterion

BT Brightness Temperature

CCDC Continuous Change Detection and Classification

COLD Continuous Monitoring of Land Disturbance

CUSUM Cumulative Sum

CWT Continuous Wavelet Transform

DAG Directed Acyclic Graph

DI Disturbance Index

DOF Degrees of Freedom

DOY Day of Year

DWT Discrete Wavelet Transform

EO Earth Observation

EOS End of Season

EVI Enhanced Vegetation Index

EWMA Exponetially Weighted Moving Average

EWMACD Exponentially Weighted Moving Average Change Detection

FI Forestness Index

GEE Google Earth Engine

GIMMS Global Inventory Modelling and Mapping Studies

GMM Gaussian Mixture Model

GMW Global Mangrove Watch

GPU Graphical Processing Unit

IC Information Criterion

IDL Interactive Data Language

IFZ Integrated Forest Z-score

LandTrendr Landsat-based Detection of Trends in Disturbance and Recovery

LASSO Least Absolute Shrinkage and Selection Operator

LIDAR Light Detection and Ranging

LST Land Surface Temperature

LULC Land use and land cover

MATLAB Matrix Laboratory

ML Maximum Likelihood

MNDWI Modified Normalized Difference Water Index

MODIS Moderate Resolution Imaging Spectroradiometer

MOSUM Moving Sum

MRA Multi-resolution Analysis

NBR Normalised Burn Ratio

NDMI Normalised Difference Moisture Index

NDVI Normalised Difference Vegetation Index

NFT National Forest Trend

NIR Near Infrared

OLS Ordinary Least Squares

PALSAR Phased Array type L-band Synthetic Aperture Radar

PDF Probability Density Funtion

PPF Percent Point Function

RLS Robust Least Squares

RMSE Root Mean Square Error

RS Remote Sensing

SAR Synthetic Aperture Radar

SOS Start of Season

SWIR Short Wave Infrared

TC Tasselled Cap

TCG Tasselled Cap Greenness

TCW Tasselled Cap Wetness

TM Thematic Mapper

VCF Vegetation Continuous Field

VI Vegetation Index

WDRVI Wide Dynamic Range Vegetation Index

YATSM Yet Another Time Series Model

Chapter 1

Introduction

Anthropogenic climate change is affecting the hydrological and ecological processes of the earth, leading to melting of the polar ice caps, rising sea levels, ocean warming, and an increase in extreme weather events [126]. While land cover can be natural or anthropogenic, population growth places increasing demands on land to produce ever greater quantities of food, materials, and shelter. As land cover changes from forest to farmland, from farmland to housing, the surface properties of that land and its interactions with the earth's processes are altered. Land use and land cover (LULC) change impacts photosynthetic activity, transpiration, and albedo, contributing to global climate change through influence on the sequestration of carbon, the cooling and warming affects of the hydrological cycle, and the reflectance of heat. Whilst not all land use changes have a negative impact, it has been suggested that agriculture, forestry and other land use change could account for 21% of anthropogenic greenhouse gas emissions [290]. Van der Werf *et al.* [310] estimated that 6-17% of anthropogenic CO_2 emissions could result from deforestation alone. It is clear that the ability to accurately monitor LULC change can contribute to understanding and mitigating the effects of climate change.

In the past few decades, Earth Observation (EO) satellites have provided an opportunity to study the earth's processes at global scale and have become an integral part of monitoring, classifying, and understanding global land use change. The im-

mense potential of satellite missions for land cover monitoring was recognised as early as 1968, when R. Colwell, founder of the International Society for Photogrammetry, stated that he could "...see the possibility that the techniques for remote sensing will evolve into a highly automatic operation, in which an unmanned satellite orbiting the Earth will carry multiband sensing equipment" [51]. Since the launch of Landsat 1 in 1972, around 200 land imaging satellites have been placed in orbit, nearly half of which were still in operation at the end of 2013 [27]. The autonomous and repetitive nature of satellite data collection results in a breadth and depth of data availability which cannot be replicated on the ground. Repeated measurements at consistent time frames allow for detailed examinations of time series data, revealing temporal trends that span years and even decades. The Landsat missions, for example, have captured imaged of the earth's surface continuously since 1972 [258]. As a result, data from satellite instruments has been widely used to monitor forest disturbance [50, 64, 250, 278, 312, 343], floods [190, 251, 332], fires [31, 101, 166, 167], and droughts [105, 268, 297], amongst many other applications.

1.1 Land cover change definitions

The Earth's surface undergoes many changes which can be detected using satellite data. These can be broadly placed into three categories as defined by Verbesselt *et al.* [297]:

- **Abrupt/step changes.** These are changes such as drought, deforestation, or construction events which involve a clear change from one type of land cover or use to another. Abrupt changes occur over short time frames (days, weeks, or months).

- **Gradual/condition changes.** These are changes in the condition of land cover due to factors such as disease, infestation, and land degradation and

recovery. In this case the underlying land use or land cover type is not changing.

- **Phenological changes.** Phenology describes the seasonal cycle which plants and animals undergo, e.g. leaf emergence, flowering, and migration. This cycle is therefore highly prevalent in vegetative land cover types. Phenological changes are changes in the timing of these events; for example, warmer spring conditions due to climate change causing plants to flower earlier [302].

1.2 Motivation

1.2.1 Data and computing availability

The opening of the Landsat archive in 2008 has provided researchers with over 40 years of continuous EO monitoring data [319], with the current Landsat 8 and 9 missions assuring that this record will continue into the near future [258]. In response, many change detection methods have been developed to take advantage of the abundance and high spatial resolution of Landsat data [338]. In addition, the launch of new high-spatial, high-temporal resolution satellites such as the Sentinel-2 missions will continue to increase demand for methods which can detect small-scale variations in time series as well as overall trends, which can operate in real time, and which can be applied regardless of geographical scale. As the cost of computing infrastructure continues to drop, this type of high throughput, live change detection becomes more and more viable. Platforms such as Google Earth Engine are also making processed satellite data more easily available and accessible without expert knowledge [19]. As a result, there will inevitably be a need for methods which are accessible in terms of software and documentation, do not require specialist knowledge to inform input parameters, can easily be integrated with other software infrastructure, and have demonstrated applicability to a wide variety of land use change types and geographical areas.

1.2.2 Advantages of time series modelling

Many change detection methods rely on either detecting whether any one observation exceeds a predefined threshold (e.g. [54, 102, 118, 122, 232, 275]) or on classifying observations then differencing between them (e.g. [31, 96, 238, 246, 326]). Thresholding methods are straightforward to implement and several specialised indices have been developed for threshold change detection such as the Integrated Forest Z-score (IFZ) [122] and Disturbance Index (DI) [232]. However, choosing the threshold value requires expertise and it may need to be adjusted based on the type of change being detected and the area being studied. Unless the threshold is adjusted to account for seasonal effects such as phenology and sun angle, care must be taken to select images taken at the same time of year. This limits the temporal and spatial accuracy of thresholding in regions with high cloud cover. Atmospheric effects present another source of error and must be corrected for.

As an alternative, classification then differencing approaches can provide information on the type of change taking place in addition to the date of change. Because each image is classified separately, errors from atmospheric effects are less likely; however, classification will be influenced by season. For multi-year vegetation change studies, images from spring and summer are often used [238, 246, 326] since this is when plant life is most easily distinguished from other land cover types and it is easier to distinguish between different plant species. As with thresholding, this requirement reduces the number of potential viable images and change detection might only be possible on a longer than yearly timescale. Another disadvantage is that the classification step introduces additional error into the change detection process. Both methods are also limited to detecting abrupt changes.

Statistical models provide representations of reality which can be used to derive information about a process which occurs over time, predict future values in the series, incorporate uncertainty about the process, and/or interpolate between values.

Rather than considering each observation in isolation or being limited to comparisons between pairs of observations, such models can take account of overall temporal context. Modelling approaches therefore possess some advantages over thresholding and classification/differencing methods. Models can act to smooth data, shifting the focus away from individual observations and towards longer-term trends and patterns. This means there is potential for gradual condition changes to be detected as well as abrupt changes, and for specific features to be extracted such as the shapes of seasonal curves.

This smoothing effect can also reduce the influence of outliers and noise, allowing for more observations to be included in the analysis and improving temporal accuracy. Statistical models can also include seasonal components which account for plant phenology, eliminating the need for images taken at the same time of year and increasing temporal accuracy to potentially within one observation. Thresholds or parameters used for change detection can also be more objective because they can be based on features of the model rather than trial and error or subjective opinion. Models can be used for interpolation, to determine missing values, and in the case of Bayesian models they can incorporate the uncertainty inherent in modelling complex time series.

1.3 Aims and objectives

1.3.1 Aims

The aims of this project are to develop a method for automated, large-scale monitoring of both land cover change and land condition change over time, and to demonstrate the usefulness of that method for a range of land cover change scenarios.

1.3.2 Objectives

The objectives of this project are:

- To analyse and document current methods for land cover, land condition, and phenological change analysis, through examination of the literature;
- To evaluate current statistical modelling techniques for analysing LULC change in remote sensing time series, including time series with incomplete or noisy data;
- To determine the applicability of dense time series monitoring approaches to large-scale forest analysis by applying a state-of-the-art change detection technique to a specific use-case of mangrove forest change;
- To demonstrate the feasibility of utilising dense time-series analysis for monitoring global mangrove forest change.

1.4 Thesis outline

- **Chapter 1** defines land cover change types and explains the motivation for the thesis.
- **Chapter 2** provides background on current methods for land cover monitoring using temporal modelling and on mangrove forest change as a use case.
- **Chapter 3** investigates the strengths and weaknesses of four state-of-the-art methods for monitoring land use change.
- **Chapter 4** applies a land cover change algorithm to monitoring long-term change in the Sundarbans mangrove forest.
- **Chapter 5** updates the algorithm used in Chapter 4 and applies it to long term monitoring of five study sites across the globe.
- **Chapter 6** provides a summary of the results and a discussion of possible future work.

Chapter 2

Background

2.1 Statistical modelling for change detection in satellite image time series

This section will detail the main temporal modelling approaches currently used for pixel-level change detection in satellite image time series. Modelling satellite image data presents several challenges. While most cloud, cloud shadow, and snow effects can be screened out using processes such as Fmask [341], it can never be guaranteed that some outliers do not remain. Any remaining effects contribute to the level of noise in the data, which is also influenced by errors from the sensor itself in addition to atmospheric effects and errors in georeferencing, orthorectification, and registration between images [173]. In addition, the need to remove noisy observations results in gaps in the time series. This precludes the use of models which require even time periods between observations, such as Autoregressive Integrated Moving Average (ARIMA) models.

2.1.1 Inter-year change detection

Inter-year change detection methods aim to detect changes at yearly temporal resolution. These include methods which aim to detect only abrupt or gradual changes, and methods which only aim to detect phenological changes, which by their nature require year-to-year comparisons between seasonal cycles. The inter-year methods

discussed here are unlikely to be useful for monitoring change at higher frequencies and have not been demonstrated capable of doing so.

2.1.1.1 Single-point

Single-point inter-year methods are those which summarize the data for each year into a single observation. Therefore, they do not attempt to incorporate seasonal cycles into time series models, instead relying on selecting images from the same time of year to discount the effects of phenology. It has been argued that, for areas of the globe affected by persistent cloud or snow cover, this type of change detection is the only option for high spatial resolution analysis [150].

Linear trend Linear trend change estimation is one of the most straightforward change detection techniques. Fitting a bivariate linear regression model to a dataset allows for the relationship between dependent (e.g. Red band surface reflectance) and independent (e.g. time) variables to be defined mathematically. A simple linear model follows Eq. 2.1, where y is the dependent variable, x is the independent variable, and a is the intercept. b is the slope and quantifies how much y changes in response to x. Whether the slope is positive or negative therefore indicates whether a change in x results in y increasing or decreasing. A test of significance is often used with linear models to either accept or reject the null hypothesis that a change in x has no effect on y (i.e. $b = 0$), resulting in a p-value. If $p < 0.05$, the null hypothesis is rejected.

$$y = a + bx \tag{2.1}$$

Vogelmann *et al.* [300] used linear regression and Landsat data to quantify changes in vegetation for four different sites in the US over two decades. Both Normalised Difference Vegetation Index (NDVI) and Short Wave Infrared (SWIR) / Near Infrared (NIR) were used depending on region and the type of vegetation being monitored. The regression model is fit to the entire time series for a set of pixels, resulting in

a slope and p-value which are used to determine whether any significant increasing or decreasing trends exist. The overall percentage of decreasing/increasing pixels is then calculated for each land cover type. Bonney *et al.* [30] used linear regression to find significant ($p < 0.05$) positive and negative trends over 30 years in Canada, using Landsat-derived NDVI. Gao *et al.* [86] use the R^2 and slope values of linear models to identify forest cover change using Moderate Resolution Imaging Spectroradiometer (MODIS) Vegetation Continuous Field (VCF) data.

$$I_w = 512 - (Band\_3 + Band\_5) \tag{2.2}$$

A similar approach, called National Forest Trend (NFT), was taken by Lehmann *et al.* [164] to look for thinning/thickening trends in forest canopies over the Australian continent. This study uses a "woodiness index" I_w rather than NDVI because of the lack of greenness in Australian vegetation. The index is defined in Eq. 2.2, where the bands used are Landsat Thematic Mapper (TM) bands three and five and 512 is an offset to make the value positive. In addition to fitting a simple linear model, Lehmann *et al.* also fit a quadratic model. Quadratic models include a quadratic term, as shown in Eq. 2.3, allowing for a curve to fitted to the data. The addition of the quadratic model allows for the detection of pixels disturbance and regrowth dynamics in addition to overall loss or gain in vegetation; for example, a positive quadratic coefficient indicates that disturbance and regrowth is likely to have occurred even if the linear coefficient suggests no overall net gain/loss in vegetation [164]. Lehmann *et al.* present a variety of sites as case studies, including areas affected by wildfire, logging, and mining.

$$y = a + bx + bx^2 \tag{2.3}$$

This type of model is a computationally fast and simple method for quantifying overall trends in time series, which has been demonstrated to be useful in detecting

vegetation changes in a wide variety of scenarios and at large scales. The downside of using a simple linear or quadratic model is that the model can only capture very broad trends, with no indication of smaller variations. Such models are also easily skewed by outliers, and highly dependent on the ability of the index used to accurately and consistently represent the underlying changes in vegetation [164].

LandTrendr Landsat-based Detection of Trends in Disturbance and Recovery (LandTrendr) uses an iterative segmentation process to find inter-year trends in Landsat time series [150]. While the Landsat missions provide a high spatial resolution, their revisit time of 8 to 16 days can easily be reduced to very few observations per year in some areas [144]. LandTrendr is based around finding points of abrupt change, called *vertices*, which separate two distinct periods in time. Each vertex is a single observation in the time series, i.e. a point with a time value and a spectral value. The gaps between each pair of vertices are called *segments*. LandTrendr can be used with any time series data [150].

The LandTrendr algorithm is an iterative process which follows a top-down approach whereby a linear model is first fitted to the entire time series, then iteratively broken up into multiple models based on the points of greatest deviation. This allows for abrupt changes to be detected in addition to gradual changes. The algorithm is carried out as follows [150]:

1. Outliers are removed from the time series by detecting observations which differ in value from the observations on either side. The degree of difference required is controlled by the *despike* parameter.

2. The first set of vertices are the first and last observations in the time series. A regression model is fitted between these points. The observation with the greatest deviation from this model is selected as the next vertex. This divides the time series into two segments. New regression models are fitted to each of

these segments, and the Mean Square Error is calculated for each model. The segment with the highest MSE is selected, and then split again by the observation with the greatest deviation. This process continues until until the number of segments is equal to the sum of two user-defined variables, *max_segments* and *vertexcounterovershoot*.

3. A culling process is carried out to remove any excess segments. The vertex in the time series with the shallowest angle is removed and a new segment is fitted between the two vertices adjacent to the removed vertex. This is repeated until the number of segments is equal to *max_segments*.

4. A second set of algorithms is used to adjust each vertex in order to find the best fitting trajectory through the time series given the vertices found. Essentially, rather than optimising the location of the vertices in time, this step optimises the spectral values of the vertices.

5. The trajectory is iteratively simplified and refitted. The purpose of this step is to determine whether a simpler trajectory (i.e. a lower number of segments) would have a better fit to the data. At each iteration, the weakest vertex is removed. The decision of which vertex to remove is either based on MSE, or on a user defined *recovery_threshold* parameter. The fitting process from Step 4 is then repeated. This results in a set of trajectories with the number of vertices ranging from *max_segments* to one.

6. The best model from Step 5 is identified using the p-values of the models.

7. For the best model, the type of each segment (disturbance or recovery) and the year and spectral values of its vertices are output. These represent gradual and abrupt changes respectively.

An additional filtering phase, which uses ground truth vegetation cover data to remove segments with a low percentage of land cover change, is also described by Kennedy *et al.* [150]. However, as this step is optional and study specific it will not be described here in detail.

LandTrendr offers a highly customisable algorithm which is available freely online [151]. In 2018, LandTrendr was also made available on Google Earth Engine (GEE) [154]. It has mainly been used for mapping long-term historical forest disturbances [76, 77, 107, 108, 174, 229]. Yang *et al.* [323] used LandTrendr to monitor land recovery from surface mining. The most common indices used are those derived from Tasselled Cap (TC) transformations [77, 107, 108, 150, 229] and Normalised Burn Ratio (NBR) [152, 153, 171, 195, 270], with some using NDVI [76, 323]. Comparing NDVI, Tasselled Cap Wetness (TCW) and NBR, Kennedy *et al.* [150] concluded that NBR was the most sensitive to disturbance events, but was slightly less accurate than TCW due to high noise sensitivity. Schneibel *et al.* [265] compared Enhanced Vegetation Index (EVI), NDVI, Tasselled Cap Greenness (TCG), and NBR, concluding that NBR was the most accurate for their study of tropical forest disturbance in Angola. While LandTrendr is a univariate method, Cohen *et al.* [50] demonstrated the potential for an ensemble approach using LandTrendr, whereby the algorithm was run multiple times, each time using a different spectral band or index. This approach has the potential to monitor multiple types of land use change as well as reducing errors [50]. While LandTrendr was developed for Landsat data, there is no fundamental reason why it could not be used for data from other platforms.

Kennedy *et al.* [150] highlight that LandTrendr is a completely relative algorithm; changes are only changes in comparison to the data around them. This can lead to many small-magnitude false changes being detected, and results in less accurate change detection for the first and last segments [150]. LandTrendr also has many user-defined parameters, which could make optimisation difficult, and the algorithm overall

is computationally expensive with several iterative steps. In addition, LandTrendr cannot detect real-time changes, and images from the same time of year must be used to minimise phenological effects.

Composite2Change The Composite2Change approach is similar to LandTrendr; however, instead of beginning with one single segment spanning the entire time series and iteratively creating new segments, Composite2Change uses a bottom-up approach which begins by fitting as many segments as possible and then merging adjacent segments [114]. The name of the method is the result of a pre-processing step whereby Best Available Pixel (BAP) composites are created for each year in the study. The BAP compositing method involves selecting the highest quality pixel for each year, based on factors such as target Day of Year (DOY), proximity to cloud and cloud shadow, opacity, and sensor [311]. The result is a set of high quality yearly images, based around a specific DOY. The selection of the target DOY will vary depending on the purpose of the study. This is in contrast to LandTrendr, which only requires that images be selected from similar times of year [150].

The change detection algorithm used by Composite2Change is as follows [114]:

1. A segment is fitted between each pair of observations so that for n observations, there are $n - 1$ segments.

2. The cost of merging each pair of adjacent segments, based on RMSE, is calculated.

3. The pair of segments with the lowest cost is merged.

4. Steps 2 and 3 are repeated until the stopping criteria are met.

The stopping criteria for Composite2Change are two user-defined parameters: Maximum number of segments, and maximum merging cost. The maximum merging

cost was included in order to avoid the issue raised by LandTrendr, whereby all changes were only defined relative to the rest of the time series [150]. The maximum merging cost can be varied during the change detection process in order to reach the maximum number of segments [114]. Once the process has finished, the overall time series trend (consisting of multiple segments) can be classified into one of four categories: No breakpoints, multiple break points all of which have positive slopes, one breakpoint with a negative slope, and multiple breakpoints with at least one negative slope. The overall trends, along with the segments and breakpoints, are used to determine metrics such as the persistence, rate, and magnitude of both abrupt change events and pre- and post-change segments [114].

Finally, a spatial analysis is carried out on the change events to detect unreliable changes [114]. Changes with less than 50% of data missing in the spatial domain for the years before, during, and after the change event are classed as reliable. Changes with more than 50% of data missing are less reliable, and if a high-reliability change exists within ± 1 year of a less reliable change, the less reliable change is re-labelled with the year of the high reliability change.

All studies found using Composite2Change were investigating forest disturbance [5, 114, 115, 137, 289, 312, 313]. Almost all of these studies were based in Canada, though one study was based in the USA [289] and one in Finland [313]. The Composite2Change approach is less computationally expensive than LandTrendr and White *et al.* [312] demonstrated that it could perform with high accuracy over large areas (the whole of Canada). Jarron *et al.* [137] found that Composite2Change could accurately detect changes in forestry caused by three different harvesting methods (clear cut, residual cut, and partial cut). All studies used NBR, in line with the original method [114]. NBR was originally chosen because Kennedy *et al.* [150] suggested that it was more sensitive to small disturbance events than other indices. All studies use Landsat data, though as with LandTrendr, it may be possible to use Compos-

ite2Change with other platforms.

The Composite2Change approach has an advantage over LandTrendr in that by using BAP composites, it is much easier to create images which are free from seasonal effects. However, it is also limited to a single index or spectral band, and so far the method has only been applied to forest disturbance. Despite these limitations, for areas such as Canada which have high levels of cloud cover which reduce observation density, Composite2Change does provide an efficient method for long term land cover change studies.

Shape fitting Unlike approaches such as LandTrendr and Composite2Change, which focus on finding the trend within the time series, shape fitting methods begin with a set of specific temporal trajectories and attempt to find the trajectory which best describes the underlying data. While no single method in this area is widely used, several different approaches appear in the literature.

In a 2007 study, Kennedy *et al.* [149] used Landsat band 5 reflectance for characterizing forest disturbance. Band 5 reflectance is high for bare ground, but low for mature forests [149]. Four potential change types are proposed: *Disturbance*, *disturbance and revegetation*, *revegetation*, and *revegetation to stable state*. *Disturbance* type changes result in step-like trajectories, where a stable low reflectance signal experiences a disturbance period resulting in a stable high reflectance period. Such changes are parametrised by three values: p_0, the end year of the disturbance period, p_1, the pre-disturbance mean reflectance, and p_2, the post-disturbance mean reflectance.

Two of these parameters, p_0 and p_1, are also used to parametrise *disturbance and revegetation* type changes. However, for this change type it is assumed that vegetation regrowth will result an exponentially decreasing curve, parametrised by the function in Eq. 2.4, where r_{b5} is the observed reflectance value, t is the time since disturbance,

p_3 is an exponential decay constant, p_2 is the mean reflectance immediately after the disturbance (the upper bound of the curve), and p_4 represents the value of reflectance as the curve approaches infinity (the lower bound of the curve) [149]. The higher the value of p_3, the faster the vegetation is recovering.

$$r_{b5} = \left((p_2 - p_4) \times e^{-p_3^t}\right) + p_4 \tag{2.4}$$

The third change type, *revegetation*, represents pixels where disturbance occurred before the start of the time series. In this case, there is no disturbance event, but there is an exponential decrease in reflectance. Eq. 2.4 is also used to describe revegetation, but with t instead being time since the start of the time series, and p_2 being the reflectance value for year one [149]. Time series where revegetation stabilizes are represented by the *revegetation to stable state* change type. This change type is the same as *revegetation*, except that an additional parameter, the year in which stable reflectance is achieved, is added [149].

All four models are fitted to the reflectance time series using non-linear least squares [149]. Analysis of Variance (ANOVA) is used to compare predicted model values with real observations and obtain a p-value for each model. The model with the lowest p-value wins. If no model achieves $p<0.05$ for the time series, the null hypothesis that there is no change is accepted.

Kennedy *et al.* [149] describe the method as being computationally expensive, however, computing power has increased substantially since 2007. While the whole algorithm is not available as a standalone package, the shape fitting software used, MPFIT, is available at [187]. Because the models are fitted to the entire time series, this method is less complex than the iterative approaches adopted by LandTrendr and Composite2Change, and well suited to studies investigating long-term trends in vegetation disturbance and recovery. The set of trajectory types can also be altered depending on the application area. Gillanders *et al.* [93] used an adaptation of

Kennedy *et al.* [149], with five different change trajectories rather than four to represent change types related to mining activity. TCG and TCW are used rather than band 5 reflectance. Similarly, the approach followed by Xue *et al.* [320] used NDVI and Modified Normalized Difference Water Index (MNDWI) with three trajectory types to characterise land use change from vegetation or water to urban.

The approach followed by Moisen *et al.* [202] is similar to that adopted by Kennedy *et al.* [149] and uses seven trajectory types to represent different change dynamics: *Flat, decreasing, jump, double jump, vee, inverted vee* and *increasing.* However, instead of fitting a non-linear model, Moisen *et al.* use a quadratic regression spline to model the time series as a smooth curve. This is a form of piecewise linear regression. The spline function used by Moisen *et al.* is defined by Eq. 2.5, where $\delta_j(t)$ are basis functions and there is a set of knots $\xi_1, ..., \xi_k$ with m being defined as $k + 1$. Based on work by Meyer *et al.* [198], the coefficients $b_1, ..., b_m$ can be constrained in order to produce the required model shape (*flat, vee* etc.).

$$f(t) = \sum_{j=1}^{m} b_j \delta_j(t) \tag{2.5}$$

For each time series all seven models are fitted and an Information Criterion (IC) is used to select the best model [202]. The model can then be used to determine a number of properties including the timing, duration and magnitude of changes. Moisen *et al.* used Landsat time series and three different indices in order to improve detection for a wide variety of forest disturbance types; these were Forestness Index (FI), NBR, and NDVI. Band 5 surface reflectance was also included. As with the other algorithms discussed, the method is not limited to a single type of index and has potential to be used for multiple different change types. In addition, the algorithm has been run over the entire conterminous US [202], proving its capability for large scale forest monitoring. Schroeder *et al.* [266] used ShapeSelectForest with the same indices as Moisen *et al.* for detecting a wide range of forest disturbance types for ten

locations in the US. An advantage over the method used by Kennedy *et al.* is that this method is much more accessible, being available as a complete algorithm for R as the ShapeSelectForest package [199].

The method used by Jamali *et al.* [135] simply involves fitting a linear, quadratic, or cubic polynomial over the entire time series. The method begins by fitting the most complex model, the cubic model ($ax^3 + bx^2 + cx + d$). The fit of the model is evaluated through testing the statistical significance of the a component using a t-test, and by determining that both a local maximum and minimum are present in the model. If the model passes both of these tests, a first order polynomial ($ax + b$) is fitted. If the second model is also found to be statistically significant the time series is assigned to the *cubic trend* class. If it is not, the *concealed cubic trend* class is assigned. The utility of the concealed trend class is that it captures pixels which have experienced increasing and decreasing values, but have no net gain or loss over the entire time series [135].

If the tests for a cubic trend fail, then the process is repeated for the quadratic model ($ax^2 + bx + c$). If the pixel cannot be assigned to either the *quadratic trend* or *concealed quadratic trend* classes, the fit of a linear model alone is evaluated. If the linear model is statistically significant, the pixel is assigned to the *linear trend* class; otherwise it is assigned to the *no trend* class. From the fitted models, it is possible to determine whether a decreasing or increasing trend is present [135]. While Kennedy *et al.* [149] and Moisen *et al.* [202] focused on high spatial resolution Landsat data, Jamali *et al.* [135] mapped vegetation trends in North Africa using an Advanced Very High Resolution Radiometer (AVHRR) 8 km resolution dataset and annual NDVI computed by averaging the highest four NDVI values from the growing season. Because the model fitting process is less complex than that used by Kennedy *et al.* [149] and Moisen *et al.* [202], this method is well suited to large geographical scales. The disadvantage is that the method used by Jamali *et al.* [135] can only find overall

trends, whereas the other methods can detect abrupt changes.

While some methods can detect both abrupt and gradual change [149, 202], the major limitation of all shape fitting algorithms is that the types of change are predefined by the initial hypothesized trajectories. Changes which do not fit these models will not be detected. While the trajectory types can be varied, the need to adapt shape fitting methods to individual land use change scenarios mean they are less likely to be widely adopted than more general methods such as Continuous Change Detection and Classification (CCDC). Shape fitting methods are also designed for detecting long-term trends and are not useful for real-time monitoring. In addition, none of the methods discussed here can take account of seasonal effects; all need yearly images selected from the same time of year.

2.1.1.2 Phenological

Phenological change detection methods aim to capture changes between seasonal cycles. While they incorporate multiple dates of observations within years, change is detected on an inter-year basis. This is done by extracting the dates of specific events in the phenological cycle, such as the start and end of the growing season, and comparing those dates between years to determine whether trends are present. Whereas single-point methods summarize yearly data in the form of maximum values or BAP composites, phenological change methods focus on smoothing in order to achieve accurate seasonal curves from which the desired data can be extracted and compared. Unlike single-point methods, phenological change detection requires high temporal resolution data in order to accurately capture seasonal curves.

Phenological change analysis involves two forms of time series modelling: One to capture the intra-year cycle, and one to capture the inter-year trend. This section is categorized by intra-year method, because the majority of methods use linear regression to find inter-year trends.

TIMESAT TIMESAT is software package providing a choice of several smoothing methods which can be used to prepare time series for phenological analysis [143]. It is available online at [146]. In addition to smoothing, TIMESAT can provide dates for the beginning and end of the growing season, the length of the season, the amplitude of the seasonal curve, and measures of seasonally active vegetation and total vegetation production, amongst others [147]. TIMESAT also allows for observation to be weighted according to their quality [147]. The data is assumed to be an index which is responsive to vegetation changes, such as NDVI.

The methods used by TIMESAT were initially proposed in 2002 by Jonsson and Eklundh [142] and further developed in a 2004 paper [143]. However, the current version of TIMESAT is available online [146] with a comprehensive user manual [147]. The most recent user manual (currently from 2017) will be used here as a reference. TIMESAT offers Savitzky-Golay, double logistic, and Gaussian-type smoothing. The TIMESAT methods for each of these will be described in their respective sections. This section provides an overview of the TIMESAT process and any stages which are shared between methods.

The first step in the TIMESAT software is to determine the number of seasons in the data. This data can be used to distinguish between single or double cropping cycles and also provides a set of local maxima and minima which are later used for smoothing. The time series is defined here as being of length N with values $I_i, ..., I_N$ and times $t_i, ..., t_N$, where N is the number of observations in the time series [147].

The model in Eq. 2.6 is fitted to the de-trended data with $\omega = 2\pi \times nyear/N$. $nyear$ is the number of years of data being used; TIMESAT requires a minimum of three years to operate, though single years can be processed by duplicating the data to create an artificial three-year time series [147]. c_1 incorporates the overall base level for the data while the two sine and cosine terms represent two seasons [147]. Fitting the model produces at least one primary maximum. If two primary maxima

are found, a threshold is used to determine whether the amplitude of the second maximum is large enough to be considered a second season [147]. The left and right minima can then be found for each season. These minima and maxima are used later for model fitting.

$$f(t) = c_1 + c_2 sin(\omega t) + c_3 cos(\omega t) + c_4 sin(2\omega t) + c_5 cos(2\omega t) \tag{2.6}$$

The three smoothing methods offered by TIMESAT are all based on Ordinary Least Squares (OLS) model fitting and aim to fit the model to the upper envelope of the time series [147]. The first method, Savitzky-Golay, is a well known smoothing method based on work by [262]. The second two TIMESAT methods are based on earlier work by the same authors in [142] and involve fitting local non-linear models to subsets of the data around the minima and maxima given by finding the number of seasons. These local models both take the form of Eq. 2.7, where c_1 and c_2 are linear parameters which determine the base level and amplitude of the curve around the minimum or maximum [147]. The non-linear parameters $x_1, ..., x_p$ govern the shape of the basis function $g(t; x)$ [147], with $g(t; x)$ being either a Gaussian type function or a double logistic function.

$$f(t) \equiv f(t; c, x) = c_1 + c_2 g(t; x) \tag{2.7}$$

The least squares approach given by Eq. 2.8 is used to optimize the parameters c and x for both the Gaussian and double logistic functions by minimizing chi^2 [147] over multiple iterations of c and x. Here the sum is over the subset of times $t_{n_1}, ..., t_{n_2}$ and values $I_{n_1}, ..., I_{n_2}$ around the maximum or minimum, where n_1 is the start of the interval and n_2 is the end.

$$\chi^2 = \sum_{i=n_1}^{n_2} [w_i(f(t_i, c, x) - I_i)]^2 \tag{2.8}$$

In order to fit the chosen model to the upper envelope of the data, the model is fit twice, with observations falling above the fitted line on the first fit being weighted more highly for the second fit [147]. The final step for both the Gaussian and double logistic methods is to extrapolate the local functions to fit the limbs of the curve as well as the maxima and minima [147]. This merging of local functions into a global function is an important feature of the method, because it allows for very flexible fitting [147]. The global function $F(t)$ is shown in Eq.2.9, where $f_L(t)$ is the function fit around the left minimum, $f_C(t)$ is the function fit around the central maximum, and $f_R(t)$ is the function fit around the right minimum [147]. $\alpha(t)$ and $\beta(t)$ drop smoothly from 1 to 0 around $(t_L + t_C)/2$ and $(t_C + t_R)/2$ respectively [147]. If multiple seasons exist in the year then the function is fitted for each peak.

$$F(t) = \left\{ \begin{array}{ll} \alpha(t) f_L(t) + [1 - \alpha(t)] f_C(t), & t_L < t < t_C \\ \beta(t) f_C(t) + [1 - \beta(t)] f_R(t), & t_C < t < t_R \end{array} \right\} \tag{2.9}$$

The ability to combine locally fitted models to extrapolate a global model fit increases the flexibility of the model, allowing it to incorporate the complexity of the seasonal curve [143].

Savitzky-Golay Eklundh *et al.* [71] and Wei *et al.* [308] both used the TIMESAT implementation of the Savitzky-Golay filter to investigate phenological changes. Savitzky-Golay filtering is based on the premise that a time series can be smoothed by replacing each value with some combination of surrounding values within a window of size n. Each value I_i is replaced by a linear combination of values from a window around I_i, calculated using Eq. 2.10 [147]. The simplest application of Eq. 2.10 filter is as a moving average, where I_i is the average of all points between $-n$ and n. In this case, the weights c_j are a constant value given by $c_j = 1/(2n + 1)$ [147].

$$\sum_{j=-n}^{n} c_j I_{i+j} \tag{2.10}$$

For a Savitzky-Golay filter, Eq. 2.10 is not used directly. Instead, a quadratic polynomial linear model is fitting to the $2n+1$ values within each window [147]. I_i is then replaced with the value of the fitted model at that time point. This method helps to preserve the widths and heights of the seasonal peaks in the data while removing noise from the signal [147]. TIMESAT can adapt the size of the window to the data, by first smoothing the data using a global value of n and then scanning the smoothed data for large increases or decreases in values around any data point i. If this is true then the window will be locally adapted to i and the smoothing carried out again [147].

Eklundh *et al.* [71] used the Savitzky-Golay method to smooth seasonal curves derived from 16-day temporal resolution MODIS NDVI data. Because NDVI can saturate in areas with high biomass, [71] also calculate an additional index, Wide Dynamic Range Vegetation Index (WDRVI). WDRVI is calculated using NDVI and a weighting value α (Eq. 2.11), which can be used to increase the dynamic range of NDVI by reducing the influence of the NIR band [71]. If $\alpha = 1$ then WDRVI is equal to NDVI. Eklundh *et al.* use $\alpha = 0.2$. All analysis was then carried out on both the NDVI and WDRVI time series [71].

$$WDRVI = \frac{(\alpha+1)NDVI + (\alpha-1)}{(\alpha-1)NDVI + (\alpha+1)} \tag{2.11}$$

The purpose of this study was to map defoliation of Scots Pine trees as a result of insect infestation, by comparing the seasonal curves of data from 2001-2002 (when little infestation occurred) to those from 2005 (high infestation) [71]. Once smoothed, the summer means and angles of the seasonal curves for the years 2000-2002 are calculated. The summer mean is calculated by taking the TIMESAT-derived seasonal midpoint, and taking a mean of the midpoint value and the two values on either side. The angle is used because insect larvae emerge around June and begin eating the pine needles, resulting in a flatter seasonal shape than for non-infested trees, which

show increasing NDVI throughout the growing season [71]. Two methods are used and compared to find the angle. The first method involves selecting two endpoints, defined as the midpoint ± 5 points, then taking the three points centred around each endpoint and fitting a linear model to them to find the slope [71]. The second method involves fitting a linear model to all of the values within ± 5 of the midpoint [71].

Both the summer means and the seasonal slopes for the three years are then averaged and two change parameters calculated for each pixel: The change in summer mean and the change in seasonal slope between 2000-2002 and 2005 [71]. A damage map is then calculated by classifying each pixel as damaged or undamaged based on a threshold [71], and the results compared to Light Detection and Ranging (LIDAR) data of the area from 2005. Generally, the correlation between the MODIS change parameters and LIDAR data was weak, possibly due to the relatively low spatial resolution of the MODIS data (250 m) [71]. In addition, the method also classifies other types of disturbance as changes, such as the clear-cutting of trees [71]. In comparing the two methods for finding seasonal angles, the second method tended to underestimate damage, while the second tended to overestimate [71]. Comparisons between NDVI and WDRVI showed no significant differences. However, the study does present a straightforward method for quantifying the rate of change of greenness during the growing season, which could be applied to any seasonal data where the midpoint of the season can be found. Application of the method to higher spatial resolution datasets could yield much higher quality results.

Wei *et al.* [308] used AVHRR NDVI data at 8 km spatial resolution to study long-term phenological change between 1982 and 2006 in China. TIMESAT with Savitzky-Golay smoothing is used to find the amount of seasonal vegetation growth for each year, defined as the integral of the area between the fitted NDVI curve and the base value for the growing season [308]. A linear regression model is then fitted to the seasonal vegetation growth values and the p-value of the models is used to

determine whether significant trends exist ($p \leq 0.05$). Significant trends can also be classed as increasing (positive slope) or decreasing (negative slope), with higher slope values indicating greater increases or decreases [308]. This method is very similar to the methods described in Section 2.1.1.1, and has similar limitations; namely, that it cannot capture very broad changes, and could easily be skewed by outliers. A more in-depth analysis could be carried out by utilising more of the phenological parameters output by TIMESAT.

Davis *et al.* [62] used TIMESAT with Savitzky-Golay smoothing to study long-term vegetation change in Namaqualand using AVHRR NDVI data. Once smoothed, Start of Season (SOS) and End of Season (EOS) are calculated using a threshold of 20% seasonal amplitude. SOS is defined as the point when the left side of the cure reaches this threshold and EOS as the point where the right side of the curve drops below it. Productivity is measured by taking integrals of NDVI over the growing season, and peak NDVI is also used as a measure of productivity. Instead of linear regression, Mann-Kendall trend tests are used to determine whether trends exist in these metrics over time. The Mann-Kendall trend test is non-parametric and produces a Kendall's τ statistic between -1 and 1 which describes the trend within the data, with 0 being no trend, a positive value indicating increasing trend, and a negative value indicating decreasing trend [62]. The Mann-Kendall significance test is used to determine the statistical significance of each trend.

Savitzky-Golay filtering is a widely used method for smoothing digital signals in areas such as medical signal processing [3, 7], chemistry [63, 169], and image analysis [12, 110]. It terms of Remote Sensing (RS) data, it can reduce the influence of noise and outliers. [32] suggest that as a local smoothing function, Savitzky-Golay has an advantage over global functions which force the data to fit a predefined model, possibly resulting in the creation of erroneous features. However, there are still likely to be small anomalies and variation in the Savitzky-Golay smoothed time series which

would be eliminated by a global function. It is also necessary to adapt the window size depending on the application, though TIMESAT does offer a solution in the form of adaptive windows.

Gaussian model Wang *et al.* [303] used TIMESAT with a Gaussian-type model to study phenological changes in the Tibetan Plateau over the period 2000-2010. Gaussian models are intuitively well suited to studying phenological cycles, which generally follow a curve similar to that of a Gaussian distribution. For this method, TIMESAT fits a Gaussian function $g(t; x_1, ..., x_5)$ which takes the form of Eq. 2.12, where $g(t; x_1, ..., x_5)$ is equal to $g(t; x)$ in Eq. 2.7 [147]. Here x_1 gives the position in time of the maximum or minimum around which the model is being fitted. If $t > x_1$ (to the right of the maximum/minimum) then x_2 and x_3 are used as they determine the width and kurtosis of the right half of the curve. If instead $t < x_1$ (to the left of the maximum/minimum), then x_4 and x_5 are used for the width and kurtosis of the left half. $x_2, ..., x_5$ are restricted in range to ensure a smooth shape [147].

$$g(t; x_1, ..., x_5) = \left\{ \begin{array}{l} exp\Big[-(\frac{t-x_1}{x_2})^{x_3} \Big], \text{ if } t > x_1 \\ exp\Big[-(\frac{x_1-t}{x_4})^{x_5} \Big], \text{ if } t < x_1 \end{array} \right\} \tag{2.12}$$

Optimal values for $x_1, ..., x_5$ are found using Eq. 2.8 and the local functions fitted around the maxima/minima are then merged to fit the entire seasonal curve using Eq. 2.9 [147]. Wang *et al.* [303] use TIMESAT to find the season start, season end, season length, and peak NDVI for each year, resulting in a set of four time series, one for each seasonal parameter, with each having 11 values. As with Davis *et al.* [62], Wang *et al.* use a Mann-Kendall trend test to analyse the relationships between the years. Mann-Kendall is used due to concerns that the derived phenological parameters are likely to be autocorrelated between years, and are therefore unlikely to fulfil the normality assumption of linear model fitting if the number of observations is low [303]. Wang *et*

al. applied the analysis to MODIS 8-day 500 m spatial resolution data. The resulting trends were compared to precipitation data, showing a high level of agreement.

Fu *et al.* [81] also studied an area of the Tibetan Plateau. They compared the Gaussian and Savitzky-Golay TIMESAT functions over several vegetation types, concluding that the Gaussian smoothing was more appropriate for that region because it retained less noise than Savitzky-Golay. However, Fu *et al.* also note that Gaussian smoothing may not be appropriate for arid and semi-arid areas with less pronounced seasonality.

The advantage of the Gaussian-type method is that it is specifically designed to be used for data containing one or more seasonal curves. It can adequately capture the individual shapes of each season within a year, whilst also having the ability to be extrapolated into a single, global model. The resulting model is less likely to be influenced by minor variations than Savitzky-Golay, and may therefore provide a more robust estimation of the underlying seasonal curve. A 2012 comparison by Song *et al.* [277] found that the TIMESAT Gaussian method produced higher correlation coefficients when correlated with the original data than the Savitzky-Golay or double logistic methods, and time series reconstructed from Gaussian fits had lower overall Root Mean Square Error (RMSE) values than either of the other two methods.

Logistic model As an alternative to Gaussian-type function fitting, TIMESAT can also fit a double logistic function around each maximum and minimum. This function is shown in Eq. 2.13 [147]. Here x_1 and x_3 determine the positions of the left and right inflection points (the points where the curve changes concavity) while x_2 and x_4 determine the left and right rates of change. As with the Gaussian function $x_1, ..., x_4$ are limited in range and are optimized using Eq. 2.8 [147]. Eq. 2.9 is used to merge the local double logistic functions into a global function [147].

$$g(t; x_1, ..., x_4) = \frac{1}{1 + exp(\frac{x_1 - t}{x_2})} - \frac{1}{1 + exp(\frac{x_3 - t}{x_4})} \tag{2.13}$$

There are several studies which compare the double logistic model with other types of model. Jonsson *et al.* [141] used both the TIMESAT double logistic and Savitzky-Golay methods to study the feasibility of using satellite data for studying the seasonal changes of coniferous forest in Sweden using MODIS data. The phenology of coniferous forests is more difficult to study than that of deciduous forest because the seasonal changes are much less pronounced [141]. In addition, in northern latitudes snow can make finding the timing of phenological events difficult [141]; for example, melting snow can be mistaken for growth. The double logistic model was found to be more sensitive to regional differences than Savitzky-Golay, and Savitzky-Golay was found to result in greater peak amplitudes [141]. However, overall the study concluded that there was little difference between the two methods.

A study by Beck *et al.* [26] compared a double logistic function with the TIMESAT Gaussian-type method and a Fourier regression method. The function used by Beck *et al.* differs from the current TIMESAT method in that it is optimised for northern latitudes where fewer observations are available in the winter. It also does not have a merging step. It should be noted that TIMESAT did not include the double logistic method at the time this study was published [143]. Beck *et al.* concluded that the double logistic function achieved a better model fit than the other two methods (lower RMSE), and that it captured the temporal trajectory at the end of the growing season more accurately than the Gaussian function. The function proposed by Beck *et al.* was also used by Julien and Sobrino [145] to study global phenological trends using Global Inventory Modelling and Mapping Studies (GIMMS) AVHRR data. Julien and Sobrino chose double logistic on the basis that it can better describe plateaus in seasonal curves where NDVI remains stable for extended periods.

Beck *et al.* built on this work in a further study ([25]), which utilised the TIMESAT double logistic function to derive seasonal parameters from MODIS NDVI data for Fennoscandia and the Kola Peninsula. These were then compared with ground observations in order to demonstrate accuracy for further studies. In this case the double logistic function was used because it was thought to most accurately represent the short growing season with rapid NDVI increases in the spring experienced by that region [25]. Zeng *et al.* [329] also used the TIMESAT double logistic method for a study of phenology across a range of northern latitudes using a combination of both MODIS and AVHRR NDVI data.

A comparison between the three TIMESAT methods by Song *et al.* [277] found that TIMESAT's double logistic function preserved the underlying shape of curves more accurately than Savitzky-Golay and on a par with Gaussian-type fitting, but fitted less well than the Gaussian method around peaks. In this case, the double logistic method was found to have a higher RMSE value than the Gaussian method [277]. Overall, it seems likely that both Gaussian and double logistic type functions are adequate tools for modelling phenological curves, with each having possible advantages depending on the data set being modelled. Double logistic functions have mainly been applied to northern latitudes and may be better suited to such areas than the Gaussian-type function.

Butt *et al.* [40] use a double logistic function to explore spatiotemporal trends in green-up and senescence in Mali, West Africa using MODIS NDVI data. Quadratic linear regression is used to find trends over both latitude and time (2000-2010). The study concludes that differences in phenology by latitude are much greater than temporal changes in phenology for the study area [40].

The method used by Zhang *et al.* [334] is similar to the TIMESAT method; however, instead of fitting a double logistic function around each maximum and minimum, it fits single logistic functions in a piecewise method. As with the TIMESAT

method, this involves finding local minima and maxima. Zhang *et al.* achieve this by finding the slope of each set of values within a moving window of 5 observations, then looking for the points where the slope changes from positive to negative or negative to positive. As with the TIMESAT method, this can result in multiple seasons being found within a year. Once these transition points are found, the logistic model in Eq. 2.14 is fitted to each section of the time series between a local minimum and local maximum, i.e. for a single period of growth or senescence [334]. Here $y(t)$ is the Vegetation Index (VI) value at time t, d is the the base VI value, and c is the maximum VI value minus d [334]. a and b are additional parameters which control the shape of the model [334].

$$y(t) = \frac{c}{1 + exp^{a+bt}} + d \tag{2.14}$$

The rates of change of these curves are then used to find key phenological transition dates, by finding local minima and maxima [334]. First, 2.15 is used to calculate the curvature K for any time t, where $z = exp^{a+bt}$, s is the length of the curve, and α is the angle of the vector tangent to the curve at time t [334].

$$K = \frac{d\alpha}{ds} = -\frac{b^2cz(1-z)(1+z)^3}{[(1+z)^4 + (bcz)^2]^{3/2}} \tag{2.15}$$

The rate of change of curvature $K^{'}$ for a time t is then calculated using Eq. 2.16 [334]. The two highest rates of change for the growth portion of the cycle are then taken as corresponding to the date of onset of leaf growth and date of onset of maximum leaf area [334]. The highest rates of change for the senescence part of the cycle correspond to date of onset of senescence and date of onset of dormancy [334].

$$K^{'} = b^3cz\left\{\frac{3z(1-z)(1+z)^3[2(1+z)^3 + b^2c^2z]}{[(1+z)^4 + (bcz)^2]^{5/2}} - \frac{(1+z)^2(1+2z-5z^2)}{[(1+z)^4 + (bcz)^2]^{3/2}}\right\} \tag{2.16}$$

In a further study Zhang *et al.* [333] applied the method to a global MODIS EVI dataset for July 2000 to March 2004. In addition to the four parameters mentioned above, Zhang *et al.* also calculate an integrated EVI value, which is the sum of the daily EVI values between the date of onset of leaf growth and date of onset of dormancy [333]. This provides a measure similar to the integrated VI value which TIMESAT can output, which indicates the overall amount of seasonal vegetation growth. A 2014 study used a similar method to study vegetation phenology in China [120]. The method was further used in a 2018 study to examine trends in phenology over 14 years using MODIS NDVI data for regions of Ethiopia [317].

It is difficult to discern any advantages in fitting two single logistic functions over fitting a single model to both halves of the growing season. Both methods assume that the slopes at either side of a seasonal peak follow a sigmoid curve. The limitations and advantages of the two methods are likely to be similar.

Fourier transform The field of Fourier analysis is concerned with decomposing time series data into a set of different frequencies, or harmonics. Each frequency describes a different periodic component of the data. A Fourier regression model incorporates these different frequencies in a linear model by including sine and cosine functions, essentially forcing the model to fit as a series of different periodic cycles. Fourier regression has often been used in RS to smooth and interpolate time series data (e.g. [34, 255, 309, 336]). The general form of a Fourier model is given by Eq. 2.17 [337], where $y(t)$ is the predicted value at time t, nf is the number of different frequencies or harmonics, f_i is the ith frequency, a_0 is the coefficient for the $0th$ frequency (i.e. the average value of the series), and a_i and b_i are coefficients for the sine and cosine components of the frequency f_i. The coefficients can be found using a method such as OLS [337].

$$y(t) = a_0 + \sum_{i=1}^{nf} a_i cos(2\pi f_i t) + b_i sin(2\pi f_i t) \tag{2.17}$$

In a 2010 study, Melendez-Pastor *et al.* [196] used the Fourier transform to study phenological changes in Mediterranean land cover types between 2000 and 2007 using MODIS EVI data. This study uses the coefficients a_i and b_i to calculate the amplitude and phase for each harmonic. A similar method is used by Breaks for Additive and Seasonal Trend (BFAST) (Section 2.1.2.1). The Pearson correlation coefficients between amplitude/phase and year are then calculated in order to study trends over time. The Pearson correlation coefficient provides a value of between 1 and -1 indicating the nature of the relationship between two variables, with 1 indicating a strong positive correlation, 0 indicating no correlation, and -1 indicating a strong negative correlation.

Anwar and Takewaka [11] used NDVI data from both AVHRR and MODIS to study phenological changes in mangrove forests in Bangladesh over a 20-year period. For this study time series are smoothed using a Fourier regression and average, maximum, and minimum NDVI are taken. Trends over time are analysed using a Mann-Kendall test.

Jeganathan *et al.* [138] used Fourier regression to smooth GIMMS NDVI at 8 km spatial and 15-day temporal resolution data in a multi-scale study of phenological change over areas above 45°N. This study highlights that there is little consensus between different studies as to how much SOS and EOS have shifted as a result of climate change, due to differences in method and number of years used in the study, and ongoing adaptation of vegetation to the changing climate. Many studies also focus on global or latitudinal trends, and do not take into account smaller scale changes, such as abrupt changes, which affect those trends [138]. This is a notable omission given that the majority of single-point change detection methods attempt

to find points of abrupt change in time series, then examine the trends between those points.

Five harmonic components were found to adequately represent the phenological curves [138]. SOS and EOS are then estimated by starting from the point of peak NDVI and comparing the difference in NDVI between pairs of observations, moving backwards and forwards in time respectively. SOS/EOS are found when the difference between two consecutive observations exceeds (for SOS) or drops below (for EOS) a user-defined threshold for four days [138]. A linear regression model is then used to find trends in SOS and EOS over time for each pixel. The mean trends are then taken for five different categories: By vegetation type, by latitudinal zone, by percentage cover loss, for core (homogeneous) pixels for each vegetation type, and for pixels with varying percentages of each vegetation type [138].

While Fourier regression is a popular method for time series smoothing, there is evidence that it results is higher RMSE values than either the TIMESAT Gaussian-type function or double logistic type functions [26]. Because Fourier regression models every point as a sum of multiple frequencies, flatter periods with no change actually require more harmonic terms to approximate [145].

2.1.2 Intra-year change detection

For areas where relatively complete, high temporal resolution data is available from platforms such as MODIS and Landsat, abrupt changes can be detected much more accurately in time if intra-year data is included in analysis. There is clearly potential for changes to be detected in a real-time manner. In addition, the successful launches of the Sentinel-2 satellites in 2015 and 2017 [73], with a combined revisit period of five days, signifies a potential step change in high spatial resolution satellite data availability across the globe. Therefore there is clearly much scope for developing methods which can detect land use and land cover change at an intra-yearly as well

as an inter-yearly scale.

2.1.2.1 Fourier regression

The application of Fourier regression to phenological change detection has been discussed. In intra-year change detection methods, the utility of Fourier regression is in incorporating seasonal information which would otherwise be flagged as change. By incorporating seasonal dynamics into a model, methods can separate the phenological cycle from long term trends and abrupt changes.

BFAST/BFAST Monitor BFAST and its variants are some of the most widely used change detection algorithms. Three forms of the BFAST method now exist; originally developed in 2009 [295], it was updated in 2010 [296] and a live monitoring version (BFAST Monitor) was developed in 2012 [297]. This study will focus on the last two methods, because the 2010 version [296] was developed as an improvement to the 2009 version [295] and uses harmonic regression rather than seasonal dummy variables to model seasonal cycles. There are substantial differences between BFAST and BFAST Monitor which will be explained here. Both methods require one outcome variable and are not specific to any one index or spectral band.

$$Y_t = T_t + S_t + e_t \qquad (t = 1, ..., n) \tag{2.18}$$

The BFAST model aims to decompose a time series into trend, seasonal, and error components [296]. This allows BFAST to monitor abrupt and gradual changes as well as changes in the phenological cycle (i.e. changes in the shape of the seasonal curve) [296]. BFAST is summarized in Eq. 2.18, where n is the number of observations, Y_t represents the observed data at time period t, T_t represents the trend component, S_t represents the seasonal component, and e_t represents the remainder component (i.e. all variation not explained by the other two terms) [296].

$$T_t = \alpha_i + \beta_i t \tag{2.19}$$

BFAST uses an iterative technique to find both trend and seasonal breakpoints in the time series. The aim of this method is to determine the number m of breakpoints with $m + 1$ segments for the trend component, and the number p of breakpoints with $p + 1$ segments for the seasonal component [296]. The number and location of breakpoints can differ between these two. The trend component T_t in Eq. 2.18 therefore depends on a piecewise linear model, with $\tau_1, .., \tau_m$ breakpoints and $m + 1$ segments each with a separate slope α and intercept β, which can be used to determine the direction and magnitude of change. This is summarised by Eq. 2.19, where $i = 1, ..., m$ and $\tau_{i-1} < t \leq \tau_i$. $\tau_0 = 0$ and $\tau_{m+1} = n$.

$$S_t = \sum_{k=1}^{K} \gamma_{j,k} sin\Big(\frac{2\pi kt}{f}\Big) + \theta_{j,k} cos\Big(\frac{2\pi kt}{f}\Big) \tag{2.20}$$

Similarly, the seasonal component S_t varies depending on the seasonal breakpoints $\tau_1, ..., \tau_p$, where $\tau_0 = 0$ and $\tau_{p+1} = n$ [296]. The seasonal component for a time period takes the form of the linear regression model shown in Eq. 2.20, where the known terms K and f are the number of harmonic terms in the model (the number of different periodicities, e.g. three, six, and twelve months) and the number of observations per year (i.e. a 16-day time series would have an f of 23) respectively. $j = 1, ..., p$ and $\tau_{j-1} < t \leq \tau_j$. The coefficients $\gamma_{j,k}$ and $\theta_{j,k}$ can be used to estimate the segment-specific amplitude $a_{j,k}$ and phase $\delta_{j,k}$ according to Eq. 2.21 and Eq. 2.22 [296]. The amplitude and phase characterize the shape of the phenological cycle, with amplitude relating to the peak of the seasonal cycle and phase relating to shifts in the timing of the cycle.

$$a_{j,k} = \sqrt{\gamma_{j,k}^2 + \theta_{j,k}^2} \tag{2.21}$$

$$\delta_{j,k} = tan^{-1}\Big(\frac{\theta_{j,k}}{\gamma_{j,k}}\Big) \tag{2.22}$$

BFAST begins by using Eq. 2.18 to estimate the seasonal component S_t [296]. The method then proceeds in an iterative manner, stopping once the number and position of the breakpoints becomes stable:

1. The seasonal component is removed from the data to obtain $Y_t - S_t$. The OLS-Moving Sum (MOSUM) test, based on work by [328], is then used to detect whether any breakpoints exist in the trend component of the model. The OLS-mosum test uses the model residuals to detect structural breaks. If breakpoints are detected, the Bayesian Information Criterion (BIC) is used to determine the number and position of the trend breakpoints, based on [21].

2. Robust regression is used to fit a linear model (Eq. 2.19) to each segment determined by the trend breakpoints $\tau_1, .., \tau_m$.

3. The trend component is removed from the data to obtain $Y_t - T_t$ and the OLS-MOSUM test is used to determine whether breakpoints exist in the seasonal component of the model. If breakpoints are detected, the BIC is used to find the number and position of the seasonal breakpoints.

4. Robust regression is used to fit a harmonic model (Eq. 2.20) to each segment determined by the seasonal breakpoints $\tau_1, .., \tau_p$.

BFAST is available as an R package [298] and has mainly been used with NDVI time series [45, 125, 140, 235, 264, 296] although EVI has also been used [231, 307]. The majority of studies used MODIS data with approximately 250 m spatial resolution [45, 125, 296, 307, 330], however de Jong *et al.* [140] used 8 km AVHRR data and Schmidt *et al.* [264] used a 30 m resolution time series created by fusing MODIS and Landsat data. One study by Platt *et al.* [235] used Landsat data only.

BFAST has also been applied at global [140], regional [45], and local scales [125] and to wide variety of locations including China [45, 123], Australia [264, 307], and the USA [41, 125]. Schmidt *et al.* [264] reported a worst-case temporal accuracy of ±40 days. BFAST has therefore been demonstrated to have wide applicability to land use change studies.

However, the use of a single index, such as NDVI, limits change detection to a specific type of land use change for each analysis. In addition, because BFAST works by segmenting the whole dataset at once, new observations cannot be analysed without re-running the entire analysis. This second limitation has lead to the development of another method by the same authors, named BFAST Monitor.

BFAST Monitor fits a model which is similar to that used by BFAST (Eq. 2.23) [297]. Here, y_t represents the data at time t, α_1 is the intercept, $\alpha_2 t$ is the slope, k is the number of harmonic terms, $\gamma_1, ..., \gamma_k$ represent the amplitudes, $\delta_1, ..., \delta_k$ represent the phases, f represents the number of observations per year, and ε_t is the error.

$$y_t = \alpha_1 + \alpha_2 t + \sum_{j=1}^{k} \gamma_j sin\left(\frac{2\pi jt}{f} + \delta_j\right) + \varepsilon_t \tag{2.23}$$

Rather than finding seasonal and trend breakpoints, and fitting a model to each segment, BFAST Monitor works by dividing the data into a stable *history period*, which is a period of time where no changes occur, and a *monitoring period*, which is analysed for deviation from the stable model. The history period is selected either using the Cumulative Sum (CUSUM) test (based on work on predicting stock market returns by Pesaran and Timmermann [228]), or by using expert knowledge of the area under study [297]. Changes are then detected in the monitoring period based on the MOSUM of the model residuals [297]. Based on both simulated and real data, Verbesselt *et al.* [297] suggest that BFAST Monitor cannot accurately detect disturbances within two or less observations of the event.

BFAST Monitor is also available as an R package [298] and has been widely used in change detection studies. As with BFAST, the dominant index used with BFAST Monitor is NDVI [65, 70, 75, 111, 297], though Normalised Difference Moisture Index (NDMI) has also been used [64, 206]. A 2015 study by DeVries *et al.* [65] only used a single order harmonic model and eliminated the trend term from the equation as it resulted in unrealistically high projected values. BFAST Monitor has been used with a wider variety of data sources than BFAST, including MODIS [297], Landsat [64, 65, 111], AVHRR [75], combined MODIS/Landsat data [70], and Landsat/Advanced Land Observation Satellite (ALOS) Phased Array type L-band Synthetic Aperture Radar (PALSAR) fused data [250]. The method used by Dutrieux *et al.* [70] is a slight adaptation to the BFAST Monitor algorithm in which an external regressor was added to improve the fit of the model; MODIS was used because it has a higher temporal frequency than Landsat. The adapted model is shown in Eq. 2.24, where the term x_t represents the observation for the external regressor at time t.

$$y_t = \alpha_1 + \alpha_2 t + \sum_{j=1}^{k} \gamma_j sin\left(\frac{2\pi jt}{f} + \delta_j\right) + \alpha_3 x_t + \varepsilon_t \tag{2.24}$$

BFAST Monitor does address the issue of real time change detection, though unlike the original BFAST it does not discern between abrupt changes and phenological changes. As a possible solution to this, Verbesselt *et al.* [297] suggest that BFAST Monitor could be used to detect whether disturbances have occurred, with BFAST being used to more closely determine the timing, magnitude, and direction of change once more data was available. Verbesselt *et al.* also suggest that the original BFAST algorithm could be used for more accurate determination of a stable history period [297]. Both of these approaches would add a substantial amount of additional processing to the analysis. In addition, a possible issue with both BFAST and BFAST Monitor is the use of OLS for model fitting. As discussed by [342], there is a risk of overfitting with OLS when using models with multiple harmonic terms. Finally,

a stable history period of at least two years is suggested [297], which could limit application in areas where multiple intra-year changes exist.

CCDC/COLD CCDC was initially developed in 2014 as a near-real time abrupt change detection method which uses multiple spectral bands for analysis [340]. Like BFAST Monitor, CCDC aims to find a stable period to which a harmonic regression model can be fitted [340]. Change is then detected through deviation from the model [340]. An updated version, Continuous Monitoring of Land Disturbance (COLD), was published in 2019 [344]. CCDC/COLD were designed for use with Landsat data, with the possibility to also include Sentinel-2 observations because Sentinel-2 collects data in similar spectral bands [340]. As well as change detection, classification of each time series segment can also be carried out. Land cover classification in CCDC/COLD is based on the assumption that different land cover classes will have different seasonal cycles and therefore different model shapes. As a result the coefficients determined in Eq. 2.25, along with other model attributes such as the RMSE, can be fed into a standard classifier given adequate training data. Typically Random Forest is used for this [66, 340, 344].

Several variants of the CCDC algorithm exist [80, 340, 342] in addition to the updated version, COLD [344]. The basis of all these variants is the model shown in Eq. 2.25, where $\hat{\rho}(i, x)$ is the predicted value for the ith Landsat band at Julian date x, $a_{0,i}$ is the coefficient for the mean of the ith Landsat band, $a_{1,i}$ and $b_{1,i}$ are coefficients representing intra-annual change, and $c_{1,i}x$ is the coefficient representing the inter-annual change, or trend [340]. T is the number of days in a year (365). The model is fitted separately to the data for each Landsat band. In [340] OLS regression was used to fit the model.

$$\hat{\rho}(i, x) = a_{0,i} + a_{1,i}cos\left(\frac{2\pi}{T}x\right) + b_{1,i}sin\left(\frac{2\pi}{T}x\right) + c_{1,i}x \tag{2.25}$$

Once the model has been fitted to the stable period, new observations are added until a change is detected. For CCDC, changed pixels are detected using the method shown in Eq. 2.26, where x is the Julian date, i is the ith Landsat band, k is the number of Landsat bands, $\rho(i,x)$ is the actual value for the ith Landsat band at Julian date x, $\hat{\rho}(i,x)$ is the predicted value for the ith Landsat band at Julian date x, and $RMSE_i$ is the Root Mean Square Error for the ith Landsat band [340].

In order to make the change detection more robust, the algorithm requires change to be flagged for three observations in a row [340]. If only one or two observations are flagged, they are discarded. If three observations are flagged, change is identified. At this point a new stable period must be found, starting from the next available observation. If an observation is not flagged as potential change, it is added to the dataset and the model is re-run to include the new data point. This adds a dynamic nature to the CCDC algorithm.

$$\frac{1}{k}\sum_{i=1}^{k}\frac{|\rho(i,x)-\hat{\rho}(i,x)|}{3\times RMSE_i} > 1 \tag{2.26}$$

CCDC involves an initialisation step to find a stable period from which change can be detected [340]. This step involves taking the next 12 available observations and fitting Eq. 4.1. Eq. 2.27 is then used to detect whether any changes occur at the start, end, or middle of the 12 observation time series. In addition to the terms already defined, here x_1 and x_n are the Julian dates for the first and last observation in the initialisation period, and T_{model} is the total time used for model initialisation. If change is detected during the initialisation period, the algorithm will remove the first observation and add one more observation until a time period is identified where no change occurs. This initialisation step is repeated whenever a breakpoint is found. The end result is a set of breakpoints and a set of models for each segment (one for each band).

$$\frac{1}{k}\sum_{i=1}^{k}\frac{|c_{1,i}(x)|}{3\times RMSE_i/t_{model}} > 1$$
$$OR \quad \frac{1}{k}\sum_{i=1}^{k}\frac{|\rho(i,x_1)-\hat{\rho}(i,x_1)|}{3\times RMSE_i} > 1 \tag{2.27}$$
$$OR \quad \frac{1}{k}\sum_{i=1}^{k}\frac{|\rho(i,x_n)-\hat{\rho}(i,x_n)|}{3\times RMSE_i} > 1$$

A 2015 study by Zhu *et al.* [342] focuses on the prediction of Landsat images from time series observations rather than change detection alone. However, because prediction is based on projecting the model forwards or backwards in time, change detection is needed in order to determine the stability of the model. Therefore, Zhu *et al.* suggest several changes to the original CCDC algorithm [342]. These include varying the number of harmonic terms in the model between two and six depending on the number of observations available [342]. This allows the model to more closely fit the data. The method also employs Least Absolute Shrinkage and Selection Operator (LASSO) regression, which uses regularization and variable selection to prevent overfitting. $2 * RMSE$ rather than $3 * RMSE$ is used in order to detect subtler changes, and six rather than three consecutive pixels are used to determine change. In addition, RMSE is temporally adjusted, to account times of year when observations may have greater variance. These changes have been incorporated by several subsequent studies [66, 223, 345].

COLD updates the original CCDC algorithm with the goal of reducing errors. To achieve this a more robust change detection method is developed by Zhu *et al.* in [344]. Firstly, in place of Eq. 2.26, Eq. 2.28 is used (notation is identical to Eq. 2.26). Rather than using a threshold of three times the RMSE of the model, COLD takes advantage of the fact that the sum of the squared model residuals follows a chi-squared (χ^2) distribution, where the number of Degrees of Freedom (DOF) is equal to the number of spectral bands. Using a Percent Point Function (PPF) with a value of 0.99, a threshold for change can be calculated [344].

$$\sum_{i=1}^{k} \left(\frac{\rho(i,x) - \hat{\rho}(i,x)}{RMSE_i} \right)^2 \chi^2(k) > \chi^2_{0.99} \tag{2.28}$$

In addition to the above, COLD uses the angle between consecutive change vectors to confirm change. The basis of this method is that a true disturbance is likely to be consistent in its direction. Zhu *et al.* suggest that a mean angle of less than 45° between consecutive change vectors indicates a persistent change, i.e. that observations after the supposed date of change continue consistently along a new trajectory (Eq. 2.29) [344]. To discount changes due to regrowth, Zhu *et al.* also propose an additional step to the change detection process which utilises the red, NIR, and SWIR bands. If NIR is increasing but red and SWIR are decreasing, land cover is getting "greener", possibly indicating regrowth rather than land disturbance. These regrowth breaks can be removed by checking whether the rate of green-up was faster or slower before the break. If it was faster, the break is likely to be due to regrowth which has stabilised [344]. Inclusion of these steps reduces commission error from both regrowth events and ephemeral change [344].

$$\frac{1}{k} \sum_{i=1}^{k-1} \beta_{i,i+1} < 45^\circ \tag{2.29}$$

Both CCDC and COLD are available from the authors as relatively easy to use MATLAB programs, but without the classification component [84]. A Python implementation of CCDC is also available as part of a set of change detection methods implemented in one package, known as Yet Another Time Series Model (YATSM) [119]. As of yet there are a limited number of studies which have used CCDC. A 2014 study demonstrated the use of both the change detection and classification elements with a use case in Boston, Massachusetts [340]. Fu and Weng [80] used CCDC as part of a larger study into the effects of urbanisation induced land use change on Land Surface Temperature (LST). In this case, CCDC was used to classify land use changes over time, producing 28 annual land cover maps [80]. These maps were then

used to compare changes in LST between different land cover types. Fu and Weng used NDVI and Brightness Temperature (BT) as inputs to the CCDC algorithm as well as the individual spectral bands, and included Landsat 8 data. Pasquarella *et al.* [223] used the YATSM version of CCDC to study insect defoliation in the US using TCG derived from Landsat surface reflectance. Another study used CCDC in conjunction with phenology metrics in order to improve classification of eight forest types [224].

In a 2016 paper, Zhu *et al.* [345] use CCDC to study greenness trends in various land use types in Guangzhou, China. As in the study by Fu and Weng [80], Zhu *et al.* include Landsat 8 data in their analysis. However, they found that Landsat 8 values were more difficult to predict than those from Landsat 5 and 7, possibly due to the different methods used for atmospheric correction. This study is also notable because after abrupt change detection and classification with CCDC, a total value for greenness change is calculated for each pixel. Because CCDC takes multiple bands into account to detect change, it results in a set of models (depending on the number of abrupt changes) for each input band. Zhu *et al.* [345] use Eqs. 2.30 and 2.31 to calculate an overall value for the start and end of each model, for each band. Here the notations for the coefficients are the same as in Eq. 4.1. i represents the ith Landsat band, and j represents the jth model. $t_{i,j}^{START}$ and $t_{i,j}^{END}$ are the start and end dates for the jth model for the ith band.

$$p_{i,j}^{START} = a_{0,i,j} + t_{i,j}^{START} \times c_{1,i,j} \tag{2.30}$$

$$p_{i,j}^{END} = a_{0,i,j} + t_{i,j}^{END} \times c_{1,i,j} \tag{2.31}$$

Eqs. 2.32 and 2.33 are then used to calculate an EVI value for the start and end of each model j [345].

$$EVI_j^{START} = 2.5 \times \frac{p_{NIR,j}^{START} - p_{red,j}^{START}}{p_{NIR,j}^{START} + 6 \times p_{red,j}^{START} - 7.5 \times p_{blue,j}^{START} + 1} \tag{2.32}$$

$$EVI_j^{END} = 2.5 \times \frac{p_{NIR,j}^{END} - p_{red,j}^{END}}{p_{NIR,j}^{END} + 6 \times p_{red,j}^{END} - 7.5 \times p_{blue,j}^{END} + 1} \tag{2.33}$$

The magnitude of gradual change in EVI over each model can then be found by differencing EVI_j^{START} and EVI_j^{END}. Eq. 2.34 is then used to sum the gradual changes over all models in order to find an overall figure for gradual change over the time series [345]. Here k is the total number of models.

$$Gradual = \sum_{j=1}^{k} (EVI_j^{END} - EVI_j^{START}) \tag{2.34}$$

The overall abrupt change in EVI over the time series is calculated by taking the sums of the differences between the estimated EVI at the end of each model and the estimated EVI at the start of the next model, as shown in Eq. 2.35.

$$Abrupt = \sum_{j=1}^{k-1} (EVI_{j+1}^{START} - EVI_j^{END}) \tag{2.35}$$

Finally, the overall change in EVI for the entire time series is calculated by adding together the total amount of gradual change and the total amount of abrupt change (Eq. 2.36) [345].

$$Total(CCDC) = Gradual + Abrupt \tag{2.36}$$

Zhu *et al.* [345] argue that by taking into account the abrupt changes in the time series, a more accurate estimate of total change in greenness can be found over fitting a single model over the entire time series, as is done by inter-year methods such as [164]. By comparing the CCDC greenness method to a simpler model, Zhu *et al.* suggest that simpler models tend to overestimate changes in greenness. This

method therefore demonstrates the potential for CCDC/COLD in quantifying long-term, inter-year trends as well as detecting abrupt intra-year changes.

Pouliot and Latifovic [239] applied CCDC in a piecewise manner to detect intra-year and inter-year change in Canada. The focus of this study was on using additional data sources (in this case AVHRR and climate data) to impute missing Landsat observations in areas with high cloud cover [239]. Observations were predicted using regression modelling with Robust Least Squares (RLS). The model used to predict Landsat observations from AVHRR data is shown in Eq. 2.37, where R is the AVHRR red band data, N is the AVHRR NIR band data, and St is the AVHRR surface temperature [239]. This approach demonstrates the flexibility of CCDC for use in areas where observations are more sparse. However, this method does require more fine tuning and parameter selection than the original CCDC method.

$$\rho_i = c_{0,i} + c_{1,i}R + c_{2,i}N + c_{3,i}St + c_{4,i}x \tag{2.37}$$

CCDC/COLD has the potential to be applied to a wide variety of land use change types due to the incorporation of multiple spectral bands and/or indices in the algorithm. The inclusion of a classification step, while not required for purely change detection studies, allows for the causes of abrupt change to be identified. CCDC has also been demonstrated to be useful for both abrupt and gradual change detection. However, over larger areas, the quantity of data required to run an analysis at Landsat scale could be prohibitive. Unlike BFAST, CCDC/COLD does not attempt to discriminate between abrupt and phenological changes, and assumes that phenological cycles are constant within stable periods [340].

The user also needs to decide whether to use a simpler approach with few harmonics, or a more complex model which risks overfitting the data and which requires more computation. CCDC/COLD needs three or six observations to confirm a change, limiting its temporal accuracy, and detecting change requires a stable period of at

least 12 observations (around 6 months of data at Landsat temporal resolution). Finally, due to the iterative model fitting process and use of multiple spectral bands, CCDC/COLD is very computationally expensive to run [340, 344].

EWMACD/EDYN Statistical control charts were developed as a form of quality control in manufacturing [273] and use *control limits* to establish when a time series deviates from a stable state. Types of statistical control chart include the CUSUM chart [220] and the Exponetially Weighted Moving Average (EWMA) chart [124]. Control charts have been used for change detection in areas as diverse as medicine [204, 210, 280, 287] and farming [157, 197, 276]. A 2014 study by Brooks *et al.* utilised EWMA charts for land cover change detection in Landsat time series data [35], naming the method Exponentially Weighted Moving Average Change Detection (EWMACD). BFAST Monitor also uses a control chart, the CUSUM chart, to find a stable history period. However, because the EWMA chart is more complex than the CUSUM chart, the EWMA method is included separately here.

EWMACD has similarities to other intra-year change detection methods in that the data is initially subset into a stable period and a monitoring period, and the stable period is modelled using harmonic regression to extract the underlying season-trend model [35]. In the next step, the model is used to predict the remaining values in the monitoring period, generating a set of residuals [35]. This is done to obtain a normally distributed set of independent observations suitable for use with an EWMA chart. EWMA charts are a type of statistical control chart which take the whole history of observations into account; however, a *tuning parameter* λ is used to adjust how the time series is weighted. The closer the value of λ is to one, the less weight is given to historical data. Brooks *et al.* use $\lambda = 0.3$ based on analysis of a subset of the data, where λ was incremented from 0 to 1 [35].

For a set of observations $x_1, .., x_n$ associated with a set of times $t_1, ..., t_n$, where

t_n is always less than $t_n + 1$, the EWMA chart value z_n at time t_n is defined by Eq. 2.38 [35]. Because this function recursively includes the z_n value for the previous time point, z_n for time t_n actually includes the entire observation history. This allows for the next value, z_{n+1}, to be computed based only on the previous value, making EWMA charts highly useful for real-time monitoring of processes.

$$z_n = (1 - \lambda)z_{n-1} + \lambda x_n \tag{2.38}$$

EWMA charts use *control limits* to flag values in the time series which are outside of normal bounds. The control limits for the chart are calculated using Eq. 2.39, where μ is the mean, n is the date, L_{EWMA} is the number of standard deviations beyond which the observation should be flagged, and s is the estimated historical standard deviation of the time series [34]. Because residual values are being used, Brooks *et al.* assumes $\mu = 0$ and s a subset of the data, designated the *training period*, to calculate a suitable value for s [35]. In the resulting chart, unchanged values (i.e. those which fit the original model) will have a value of zero, with deviating pixels producing positive or negative values depending on the size of the deviation from the original model [35]. This chart can therefore be used to identify change points in the time series.

$$CL = \mu \pm L_{EWMA} s \sqrt{\left(\frac{\lambda}{2 - \lambda}\right)[1 - (1 - \lambda)^{2n}]} \tag{2.39}$$

In addition to simply flagging changes, EWMA charts can be used to determine the magnitude and direction of an abrupt change by dividing the EWMA chart value at time t with the control limit value at time t [34]. In combination with their real-time monitoring capabilities, this makes EWMA charts well suited to applications which need to quickly detect and quantify new changes. Brooks *et al.* applied the EWMA method to a time series of Landsat data using a TC derived index in order to detect three types of forest disturbance: *Light thin*, *heavy thin*, and *clearcut* [35]. The

method was typically able to detect changes within one image of their occurrence, demonstrating high temporal accuracy [34] and an improvement over CCDC/COLD, which requires at least three observations to confirm change.

A disadvantage of the EWMACD is that it assumes a single disturbance. In 2017, Brooks *et al.* published an updated version of EWMACD, called Dynamic EWMACD, or Edyn [37]. Edyn includes a re-initialisation step, whereby once a change is flagged, the data is subset to discount all data before the change and a new stable period is identified. The addition of this step makes Edyn much more similar to the BFAST and CCDC style algorithms, in that it allows Edyn to flexibly monitor for changes in the longer term. A vertex approach similar to that used by LandTrendr is used to discover when the time series stabilises after a change [37]. Brooks *et al.* found Edyn to have more omission errors and fewer commission errors than EWMACD, since Edyn is more likely to miss changes due to re-initialisation of models after breaks [37]. The authors suggest that users consider which variant will give the best results for their specific scenario(s) [37].

EWMACD/Edyn has been found by multiple studies to perform well in terms of omission/commission error when compared to other change detection methods [37, 50]. However, EWMACD/Edyn is designed specifically to be capable of detecting partially changed pixels and subtle changes in land cover condition [35, 37]. This makes it unsuitable for change detection analysis where the main goal is to detect between-class changes. So far, EWMACD/Edyn has mainly been used on single-index NDVI time series [35, 37, 50, 263]. Potential for a multi-band approach has been highlighted by Brooks *et al.* [35, 37], however, no such approach has yet been developed.

2.1.2.2 Wavelet analysis

Algorithms such as CCDC and BFAST Monitor use harmonic regression in order to account for the seasonal cycle of intra-year time series data. Such methods assume

the presence of multiple periodic cycles of different frequencies. However, they give no indication of where in time those frequencies occur; each frequency is assumed to be infinite. Wavelet analysis offers an alternative whereby both the timing and frequency of different time series components can be preserved. This is possible because wavelets have *compact support*, meaning that their value becomes 0 outside of a specific time period.

A *mother wavelet* $\psi(t)$ is used to decompose a signal by generating a set of *daughter wavelets*, each of which has a different scale, determined by the scaling parameter a, and a different location, determined by the localization parameter b [189]. By translating the mother wavelet in both space and time, a wavelet transform can be used to construct a *time-frequency* representation of a signal, whereby information on both the frequencies present and the times when they occur is preserved. This is possible because each wavelet will better fit some parts of the time series over others, highlighting areas where there is a close correlation between the wavelet and the data. The Continuous Wavelet Transform (CWT) over all scales and locations ($W(a,b)$) of a signal $f(t)$ is then given by Eq. 2.40, where ψ^* is the complex conjugate of the mother wavelet [189].

$$W(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} \psi^*_{a,b}\left(\frac{t-b}{a}\right) f(t)dt \tag{2.40}$$

However, satellite data time series are not continuous but discrete, i.e. the observations are taken at specific time steps. The Discrete Wavelet Transform (DWT) provides an approach for dealing with such signals, whereby the daughter wavelets are discretely scaled [189]. The discrete time series $f(t_i)$ can be described as a set of N observations where $i = 1, ..., N$ [189]. In order to decompose $f(t_i)$, the values $a = 2^j$ and $b = k2^j$ are used to scale and translate the mother wavelet, where j is the level of decomposition and k is a location index where $k = 1, ..., 2^{-j}N$ [189]. A discrete daughter wavelet can therefore be described by Eq. 2.41 [189]. As a and j

decrease, the resolution of the analysis increases, so that smaller components of the time series can be found.

$$\psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k) \tag{2.41}$$

In a 2009 study, Martinez and Gilbert [189] implemented a Multi-resolution Analysis (MRA) using a hierarchical algorithm as proposed by Mallat in a 1989 paper [185]. This method involves successively decomposing the time series signal into a *detail* component and an *approximation* component for M levels. This is done by using a set of high- and low-pass filters derived from the wavelet coefficients and a scaling function [189]. At each decomposition level, the detail component D is the high-pass filtered component and the approximation component A is the low-pass filtered component. The approximation component for one level is used as the basis for the next decomposition, so that $A_m = A_{m+1}D_{m+1}$. Eq. 2.42 describes how the original signal can be reconstructed from these components [189].

$$f(t) = A_M(t) + \sum_{j=1}^{M} D_j(t) \tag{2.42}$$

Martinez and Gilbert [189] applied the MRA to NDVI data for six pixels in Spain, representing several land use types including forest, rice, and cereals. Six decomposition levels were used, with the approximation component for A_6 describing inter-annual variability and the detail components $D_2, ..., D_6$ describing intra-annual variability at different scales. D_1 is discounted as being noise because it represents cycles of less than one month. A set of metrics are derived from the MRA, including the mean and slope of the inter-annual component (A_6), and the date of maximum NDVI. Sen's slope method [269] is used to determine the trend for A_6 and the Mann-Kendall test is used to quantify the significance and direction of the trend. From this the overall greening or browning trend for each pixel can be found.

The Mann-Kendall/Sen's slope method was also used by Yan *et al.* [321] to study greening/browning at the mouth of the Yangtze river using MODIS-derived VIs and by Wang *et al.* [304] for a study of north-western China using 8 km resolution AVHRR data. While both Martinez and Gilbert and Wang *et al.* used a mother wavelet called the Meyer wavelet, Yan *et al.* used the Coiflet 4 wavelet. The Meyer wavelet is claimed by both Martinez and Gilbert and Wang *et al.* to be the most suitable due to ease of implementation, low computational cost, and regularity condition.

Because the time series is decomposed into multiple scales, wavelet analysis has potential for detecting a wide variety of changes, including both abrupt and gradual change as well as shifts in phenology over time. Different wavelets have different shapes and in theory multiple analyses could be carried out on the same data with various mother wavelets, in order to detect different types of change. Lau and Weng [162] used the Morlet wavelet for MRA of long-term climate data, demonstrating its utility in finding both short oscillations and major trends. The MRA is computationally fast and preserves much more information about the time series than harmonic-based methods. There is a freely available Python library for performing wavelet transforms [163].

However, MRA has not been widely applied to change detection in satellite image time series, though wavelet analysis has been used for time series smoothing [78] and identifying the timing of phenological events [85, 261]. This is possibly because the method requires the correct choice of mother wavelet and decomposition level to be effective. Also, it does not allow for real-time change detection, and additional steps are needed in order to compute the timing, direction and magnitude of changes after the MRA has been carried out. In this sense, it is more similar to the inter-year methods discussed in Section 2.1.1.

2.1.2.3 Bayesian

Frequentist statistics measures probability as the likelihood that a certain outcome will occur given its frequency in an ensemble of trials or experiments [56]. That is, if an experiment were to be repeated many times, the outcome x would occur a certain number of times and the outcome y would occur a certain number of times. Bayesian statistics focuses instead on the "reasonable expectation" of an event occurring based on prior knowledge [56]. If more information becomes available, the prior knowledge can be updated, revising the estimate of the likelihood of the event. This is done using Bayes' theorem.

Bayes' theorem allows us to calculate the probability of an event A given that we know an event B has occurred upon which A is *conditionally dependent*; that is, we know that the probability of A is dependent on the probability of B. The probability of A given B is known as the *posterior probability* and can be calculated according to Bayes' theorem which is stated in Eq. 2.43. Here $P(A)$ represents the overall probability of A (known as the *prior probability*) and $P(B)$ represents the overall probability of B (known as the *evidence*). $P(B|A)$ is known as the *likelihood* and represents the conditional probability of B given A. The utility of Bayes' theorem is that it allows for an estimate of the probability of an event to be made (the prior probability), which can then be updated given new evidence to give the posterior probability.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}) \tag{2.43}$$

One application for Bayes' theorem is in the field of Bayesian networks. Bayesian networks are a way of representing a system as a series of variables and conditional probabilities. They are a form of Directed Acyclic Graph (DAG). An example of a DAG is shown in Figure 2.1. Here there are five variables (vertices) and three edges, which represent the conditional dependencies. The edges indicate that D is

conditionally dependent on both B and C, that B is conditionally dependent on A, and that C and B are conditionally independent.

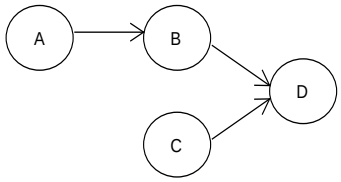


Figure 2.1: A Directed Acyclic Graph with five vertices and three edges.

Each variable has multiple possible outcomes (events); generally, these outcomes are discrete although some implementations allow for continuous variables [46]. The *probability distribution* of a variable describes the probability of each outcome. For example, let variable A have two possible outcomes a_1, a_2 with the probability distribution $0.5, 0.5$ and B have two possible outcomes b_1, b_2 with the probability distribution $0.7, 0.3$. Because B is dependent on A, a *joint probability distribution* table is used to give the probabilities of each possible combination of outcomes, that is, how likely each combination of events is to occur (Table 2.1).

Table 2.1: Prior joint probability distribution for A and B.

	a_1	a_2	
b_1	0.4	0.3	0.7
b_2	0.2	0.1	0.3
	0.6	0.4	1

This prior set of probability distributions must be determined either through expert knowledge or by using existing data to train the network. If new information is later revealed, for example, that the outcome of A is definitely a_1, then the probability distribution of A becomes $1, 0$ and the probability distribution of B must change to reflect this, since a_2 can no longer occur. Since B is a descendant of A, this is called *predictive propagation*, i.e. the new information is propagating forward through the network. It is also clear that if our new information is instead about B, the probability distribution of A must change, since some outcomes of A are more likely to lead to

some outcomes of B. This is called *retrospective propagation*. Eq. 2.44 can be used to calculate the probability distributions of each variable given new information about another variable. Here a represents the prior marginal probability of the observed event (in the example above this would be 0.6).

$$P(A, B, C, D|a) = \frac{P(A, B, C, D)}{P(a)} \tag{2.44}$$

Updating Table 2.1 according to Eq. 2.44 for a new probability distribution of $1, 0$ for A results in the joint probability distribution shown in Table 2.2.

Table 2.2: Posterior joint probability distribution for A and B.

	a_1	a_2	
b_1	0.7	0.0	0.7
b_2	0.3	0.0	0.3
	1.0	0.0	1

Bayesian inference can therefore be used to assign an outcome, such as *forest* or *not forest*, to a pixel value based on the probability of that outcome given other variables. Bayesian networks have been used for change detection in many areas including climate monitoring [259], traffic movement [148], and fault monitoring [221]. In RS, Bayesian networks have also been used for flood detection using SAR data [58, 247]. Symeonakis *et al.* [282] used Bayesian inference to improve land use classification by including temporal information; however, change detection is still carried out using post-classification comparison. Çelik [42] used a Gaussian Mixture Model (GMM) and Bayes' theorem to classify pixels as either changed or unchanged. This approach is essentially an improvement over simpler differencing techniques, since it eliminates the need to set a manual threshold.

In a 2015 study, Reiche *et al.* [248] used a Bayesian approach to fuse optical Landsat NDVI data and Synthetic Aperture Radar (SAR) data from ALOS PALSAR for monitoring deforestation, which takes account of past and future observations. Because SAR data is not affected by cloud cover and can be collected at night, fusing

SAR and optical data can result in more accurate change detection in areas where the temporal frequency of optical data alone may be low. Reiche *et al.* begin by defining n sets of time series $s1, ..., sn$ where each time series is derived from a different sensor. It is assumed that the sensors do not collect observations at the same frequencies and that each time series may have gaps [248]. Each time series has a set of observation times $T_{sn} = t^{sn}1, ..., t^{sn}k_{sn}$ where k_{sn} is the number of observations for the sensor. Finally, each observation is denoted by sn_t where $t \in T_{s1}, ..., T_{sn}$ [248].

Reiche *et al.* then used a Maximum Likelihood (ML) approach to obtain a Probability Density Funtion (PDF) for each of the F and NF classes. These are calculated separately for each sensor type. PDFs are used for continuous variables and, as for discrete variables, they map specific outcomes to specific probabilities. This allows for a probability of being NF and of being F to be calculated for each observation sn_t. Eq. 2.45 is then used to provide an overall probability of each observation being NF [248], resulting in a time series of probabilities s^{NF}.

$$P(NF|sn_t) = \frac{P(sn_t|NF)}{P(sn_t|NF) + P(sn_t|F)} \tag{2.45}$$

A flag λ for each observation is set to 1 (potentially deforested) for an observation if $s_t^{NF} > 0.5$ or to 0 (not deforested) otherwise [248]. A major element of this method is that Bayesian updating is now used to include past and future observations in determining whether an observation truly represents a deforestation event. If $\lambda = 1$ for time t, the previous observation $t-1$ and future i observations where $i = 0, ..., n$ are used to calculate the overall probability of deforestation at time D_t according to Eq. 2.46 [248]. Here $P(D_t|S_{t+i}^{NF})$ represents the posterior probability that deforestation occurred at time D_t given the new evidence S_{t+i}^{NF}.

$$P(D_t|S_{t+i}^{NF}) = \frac{P(S_{t+i}^{NF}|D)P(D_t|S_{t+i-1}^{NF})}{P(S_{t+i}^{NF})} \tag{2.46}$$

Eq. 2.46 is used iteratively to keep updating the posterior probability at time t [248]. On each iteration, the current posterior probability is updated with the new evidence from $t + i$. For $i = 0$, the probability $s^N F$ for the current time is used. The algorithm stops when a predefined threshold χ is reached ($P(D_t|S_{t+i}^{NF}) \geq \chi$). After experimenting with different values of χ, Reiche *et al.* used 0.5 for SAR observations and 0.975 for NDVI observations. The study applied the method to detecting deforestation events in Fiji on a three-monthly time scale, with reference data supplied by a logging company. They found that incorporating both optical and SAR data improved both the spatial and temporal accuracy of the change detection when compared to a single data source. Building on the 2015 study, a 2018 study [249] applied the same method for fusing Landsat, ALOS-2 PALSAR-2 and Sentinel-1 for deforestation detection in dry tropical forest in Bolivia. The fused time series improved temporal accuracy by up to 56 days in comparison to single-sensor time series.

The clear advantage of this method is the ability to easily incorporate multiple data sets into one time series change detection algorithm. The method is also able to operate in near real time, and the use of probability distributions reduces the effects of noise and seasonality. However, only abrupt changes can be detected. Defining the PDFs also requires training data and this training step, along with defining the χ threshold, are land use type specific. Nevertheless, Reiche *et al.* [248, 249] demonstrate that the method has high applicability for forest change detection in areas where clouds, cloud shadows, or snow preclude intra-year change detection with optical sensors only. Elements of the Çelik [42] method could be combined with the Reiche *et al.* method to create a none-land use type specific Bayesian network which could be applied to single-sensor data.

2.1.2.4 Artificial Neural Networks

Artificial Neural Networks (ANNs) are loosely based on the structure of biological neural networks and were proposed as early as 1943 [192]. Deep learning has since been applied to a wide range of classification and prediction problems in fields such as medical image analysis [281], language processing [219], and image recognition [82]. ANNs consist of a structure made up of *neurons* (or *nodes*) and the connections between them, known as *vertices* or *edges*. Each neuron takes one or more inputs from previous neurons and outputs a single value based on the *weights* of those inputs plus a value specific to the neuron, which is known as the *bias*. Output is further attenuated by applying a nonlinear *activation function*. The final layer is the output from the network and provides the desired prediction given the input, which could be a classification, the location of an object or objects, or a segmented image, among many other possibilities. Initially, all weights and biases are randomly initialised, and are refined over time as the network is trained. Training is carried out by "showing" the network many training examples. Each example in the training set is provided as input to the network, and the weights and biases are adjusted using an optimisation method, usually stochastic gradient descent [279]. Changes are *backpropagated* through the network until all weights and biases have been updated. Over time, as the network is shown more training examples, it learns to produce the correct output for each input. ANNs can therefore capture a wide range of nonlinear relationships.

Figure 2.2 shows the structure of a simple ANN, known as a Multilayer Perceptron (MLP). This structure forms the basis of all ANNs, including those referred to as *deep* neural networks. The structure consists of an input layer $x_1...x_n$, a hidden layer $h_1...h_n$, and an output layer $o_1...o_n$. The hidden layer is so called because its values are hidden during training; i.e. the user typically only sees the input to and output from the network. Deep learning architectures are so called because they contain multiple

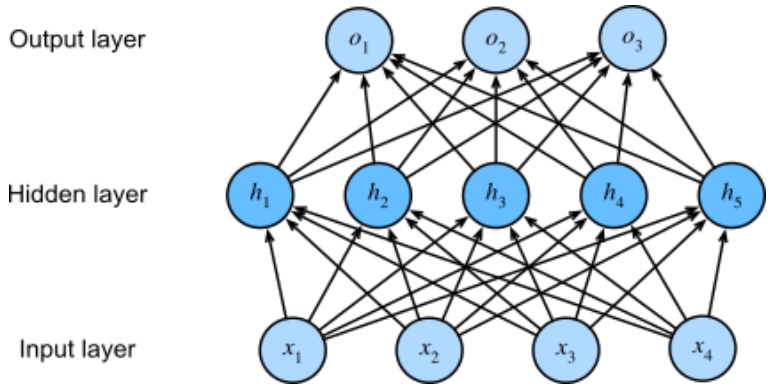


Figure 2.2: Structure of a Multilayer Perceptron. From d2l.ai [331].

hidden layers between the input layer and the output layer. The presence of these hidden layers mean that even more so than traditional machine learning techniques, deep learning models are often referred to as "black boxes", because the values in these hidden layers are difficult to interpret.

Bischof *et al.* used an ANN for classifying land cover from Landsat imagery as early as 1992 [29]; however, until the mid-2010s, very little research in the field of remote sensing was carried out which utilised deep learning models. The focus here will be on studies which use deep learning models specifically for time series modelling of satellite image time series, rather than those which aim to detect change between bi-temporal image pairs. Studies within this definition almost exclusively involve the use of Long Short-Term Memory Networks (LSTMs). LSTMs are a type of neural network which is specifically designed to operate on time series. An LSTM cell replaces the standard neuron with a more complex set of operations, shown in Figure 2.3. For an input time step X_t and output activation O_t, the input gate I_t controls how much the new data is taken into account while the the forget gate F_t controls how much of the content of the previous memory cell C_{t-1} is retained. A combination of these three values, in addition to the hidden state H_{t-1}, generates the new memory value C_i. LSTM units can therefore be used to learn sequential information, whereby the outputs from a cell for time point t are used as inputs

for the next cell in the sequence. This structure allows LSTMs to retain relevant information from previous time points.

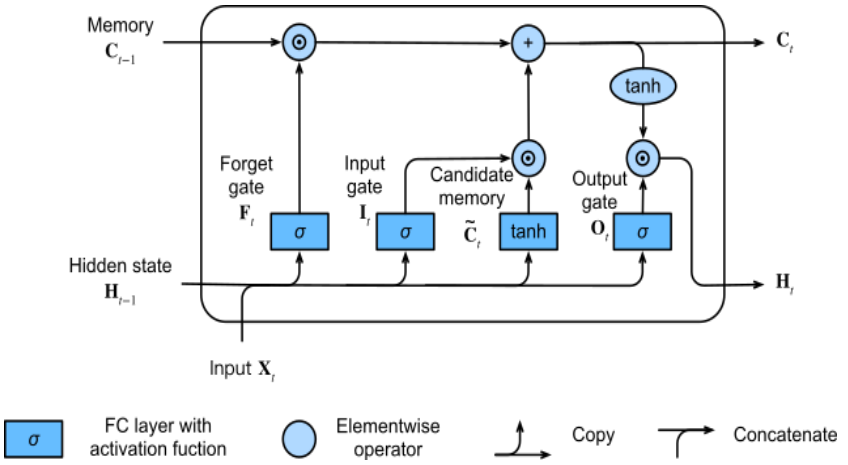


Figure 2.3: Architecture of a Long Short-Term Memory cell. From d2l.ai [331].

In remote sensing, LSTMs (and ANNs generally) have mostly been used to generate land cover classifications which take into account the phenology of the underlying vegetation (e.g. [128, 139, 170, 181, 191, 335]). While not directly useful for change detection, the generated yearly land cover maps can then be used to track changes between years. This approach is similar to modelling methods such as CCDC and BFAST which incorporate seasonal variation; however, in the case of LSTMs, modelling and classification are parts of the same model rather than a two-stage approach of model fitting then classifying based on coefficients. A disadvantage is that unlike the season-trend modelling methods, this approach does not allow for change to be detected within years. In that sense it is more similar to the phenological change approaches detailed in Section 2.1.1.2.

Current research suggests that there is potential for LSTMs to be used for time series prediction and anomaly detection. Both Rhif *et al.* [252] and Yu *et al.* [325] used LSTMs for prediction of NDVI time series. A 2015 study by Malhotra *et al.* tested the use of LSTMs for anomaly detection on four real-world datasets, concluding that LSTMs may be a viable approach for detecting deviation from normality in time

series [184]. However, the time series used in this study were much more consistent in frequency and less complex than remote sensing time series (power consumption, electrocardiograms, and engine sensors). Approaches such as Convolutional LSTM [274], developed for precipitation models, have demonstrated the potential of deep learning for prediction which takes into account both spatial and temporal information. This is highly relevant to satellite image analysis, where pixel-based approaches lack spatial context which often contains useful information.

It is clear that while there is potential for deep learning approaches to be applied to land cover monitoring, the field is still in its relative infancy. While CNNs are effective classifiers, they are no different from a monitoring perspective than other classification-then-differencing approaches which look at individual time points. This can be mitigated through the use of time series approaches such as LSTMs, but the need for large, labelled data sets is a major issue for any deep learning approach. To achieve good results requires hundreds if not thousands of samples, which in the case of satellite data often have to be human generated - a subjective and time consuming task. For example, training an effective LSTM-based network for anomaly detection would first require the generation of a substantial training data set with labelled anomalies, representing all possible change types. This is very difficult since it requires a definition of what represents an anomaly in a satellite image time series. Unlike approaches such as CCDC, where change is detected as deviation from a fitted model, LSTMs provide no error estimate and it is therefore not possible to evaluate whether a data point falls outside of the range of normal variation. In addition, LSTMs work best when presented with regular and complete time series, which is a challenge for many applications involving satellite image time series.

Deep learning models are also very computationally expensive to train, often requiring access to Graphical Processing Units (GPUs) rather than CPUs to accelerate training, and require extensive parameter tuning to achieve the best performance.

Because of the black box nature of neural networks, parameter tuning can be time consuming, often requiring a trial and error approach. Unlike traditional modelling approaches which produce coefficients for trend and seasonality, ANNs provide little insight into the nature of the relationships between variables. This means that despite the fact that they can encode long-term temporal dependencies, information on land cover condition change or other useful patterns cannot be extracted from a neural network. In summary, while this is a very active area of research, ANNs currently offer few advantages over existing land cover monitoring algorithms in terms of accuracy or interpretability, and present substantial difficulties in terms of training and implementation.

2.1.3 Summary

The field of change detection in satellite image time series is broad and many approaches have been developed. As a result it is often difficult to draw comparisons between methods. Spatial and temporal accuracy are important; however, a study by Cohen *et al.* which included both inter- and intra-year methods found that disturbance maps created from different change detection algorithms rarely agreed about either the timing or location of changes [50]. Lack of agreement between methods is likely to be because different algorithms use different spectral bands or indices, aim to detect different magnitudes of change, have different change thresholds, and use different methods for removing noise and/or outliers which could influence the data [50]. In addition, methods which aim to detect gradual changes are very difficult to validate. However, there are several factors which affect the applicability and usability of all change detection methods.

One fundamental aspect of change detection studies is the number of dependent variables included in the model. The majority of methods are designed to use a single spectral band or index, such as NDVI. For phenological change detection, it is

important that the chosen index can capture the full range of growth and dieback present at the study site(s). Given that most applications of change detection focus on a limited number of land use types, a single band/index is often sufficient. For example, a common application of land use change algorithms is to forest change, either to monitor abrupt changes from logging or gradual trends in greenness/brownness. This is not surprising given the role of forests in storing atmospheric carbon [121]. All of the methods discussed in this study have been applied to forest at least once; some, such as LandTrendr, have been used almost exclusively for forest monitoring. Such studies select the land use type of interest before running the change detection algorithm. However, much data is lost when using a single index or band. Single indices can also struggle to incorporate a wide range of values; for example NDVI can become saturated in areas with high biomass [59, 296].

For more general land use change scenarios, there is potential for less targeted analysis which incorporates more spectral information or includes more than one approach. If a land cover change algorithm is tuned to perform well for one geographical area, land cover type, or change type, it becomes less accurate when applied to other use cases. This generalization problem in land use change methods is discussed in detail by Healey *et al.* [113]. Healey *et al.* investigate the utility of combining the outputs of multiple change detection algorithms with additional data such as topography and forest type information to improve accuracy. Ensemble methods can also involve multiple applications of the same algorithm, each time with a different band/index as the input. Cohen *et al.* apply this approach with LandTrendr [50]. Alternatively, multiple variables can be tested on a subset of the data in order to determine the best single band/index for an analysis (e.g. [64]). The disadvantage of ensemble approaches is the increased computation time and complexity; however, as discussed by Healey et al., increased access to high performance cloud computing platforms such as GEE, which also facilitate code/program sharing, will likely reduce computational

complexity as a barrier [113].

CCDC/COLD proposes an entirely different approach, whereby five Landsat bands are used for change detection simultaneously. An observation is only flagged as change if it appears strongly enough in multiple bands. This method has the advantage of only needing to be run once per pixel, in addition to being applicable to a wide variety of land use types. A disadvantage of CCDC/COLD could be that the analysis loses accuracy through being less land use type specific. The Bayesian approach used by Reiche *et al.* [248] also includes multiple sources of information; however, the two data sources are analysed separately and each is only used to confirm changes found in the other.

It is likely that single-variable methods will continue to dominate; however, multivariate methods offer an attractive alternative for broader land use change studies. No phenological change detection studies attempt to combine information from multiple indices or bands, though Eklundh *et al.* do compare NDVI and WDRVI [71]. While NDVI is generally considered to be effective for vegetation monitoring, it is possible that data from other bands/indices could help to define phenological transition points more accurately. More recent studies, such as those applying deep learning, usually use multiple spectral bands for analysis.

In addition to the variable(s) used as model input, models vary widely in the number of user-defined parameters needed to perform analysis. Some methods, such as LandTrendr, have a high number of such parameters. Many methods require at least one. Composite2Change requires a user-defined stopping point and merging cost, EWMA charts require a tuning parameter, wavelet analysis requires a choice of mother wavelet, and the Bayesian approach use by Reiche *et al.* requires a threshold [248]. Phenological change detection involves selecting which phenological parameters are of interest. These values are inherent to their respective methods and their selection is an important step in the analysis. The advantage of these parameters is

that they provide flexibility, allowing the sensitivity of the analysis to be adjusted and for the methods to be optimized for different regions, land use types, or temporal resolutions. Parameters such as thresholds are fairly intuitive and can be based on previous studies, or estimated from the data. Deep learning approaches generally require extensive parameter tuning (e.g. Choice of activation function, number of hidden layers, number of epochs, learning rate) which requires reasonable knowledge of network architecture and function.

However, optimization can take time, especially if there are multiple parameters, and can also require specific knowledge of the region under study or the method itself to be effective. For example, the choice of mother wavelet for a wavelet decomposition will have a large impact on the method's outcome, and there is currently no consensus as to which wavelet is best suited to RS data. Parameters such as the λ value necessary for EWMA chart analysis or the merging cost used by Composite2Change are also less intuitive and rely on trial and error. Methods which require fewer user inputs, such as CCDC/COLD, may be more objective. Minimizing the number of user-controlled parameters and/or providing defaults which are known to work well for a wide variety of use cases could also increase the likelihood of a method being more widely adopted.

One factor which only affects intra-year change detection is whether methods can operate in near-real time, i.e. at the temporal resolution of the sensor. Methods are generally considered able to operate in near-real time if new observations can be added to the model and flagged as change/not change without re-analysing the entire time series. This ability reduces computation time for applications which need to monitor changes continually with minimal temporal delay. Both BFAST Monitor and EWMA/Edyn have the capability to detect changes within one observation, whereas CCDC/COLD and the Bayesian network approach require multiple observations to confirm a change. However, Verbesselt *et al.* suggest that the probability of detecting a true change event with BFAST is much higher if more than two new observations are

used [297]. Given that pixels may contain multiple land use types, that disturbance events could occur over days for a single pixel, and that clouds or snow could result in missing observations, the multiple consecutive observations approach may be more robust.

A downside of the EWMA, Bayesian, and LSTM methods is that they contain no trend information. Once a disturbance has been detected, CCDC and BFAST Monitor both include trend terms, allowing for condition changes to also be monitored between breaks. However, none of the live change detection methods discussed here are able to discriminate between changes caused by abrupt change and those caused by changes in phenology. Shifts in the phenological cycle, such as change in amplitude or lengthening of the season, could lead to increased RMSE or greater variance in the same way as an abrupt change in land use.

A disadvantage of most of the near-real time monitoring approaches is the need for a stable history or training period. BFAST Monitor, CCDC/COLD, and EWMA charts all detect abrupt change by looking for deviations from a stable model. If no stable period can be found then no analysis can take place. Verbesselt *et al.* suggest that BFAST Monitor requires a history period of at least two years [297]; Zhu *et al.* suggest a minimum of 12 observations covering at least one year for CCDC/COLD [340, 344]. Brooks *et al.* suggest the minimum number of training observations for EWMA/Edyn should be $3(1 + k_s + k_c)$, where k_s and k_c are the number of harmonics for the sine and cosine terms respectively. This requirement prevents the application of these algorithms in areas with high overall levels of disturbance, and precludes them from being used for analysis on data from recently launched satellites. Whilst it does not provide trend information, the Bayesian approach proposed by Reiche *et al.* does not require a history period, instead relying on predefined probabilities of a pixel being forest or non-forest based on its NDVI value [248]. Because this method also incorporates both optical and SAR data, it could be especially useful for

real-time monitoring in areas with infrequent observations.

Finally, the availability of a method is also important. A straightforward method such as simple linear modelling is possible to implement without great difficulty. However, most methods are more complex, with iterative or multi-stage processes. The most widely used methods are BFAST and BFAST Monitor, which are available as a package for the popular R programming language [298]. The popularity of these methods is probably at least partly due to their accessibility, though the fact that BFAST is the only method which can account for all three change types may also be a factor. Another example is LandTrendr, which has a well-documented implementation using a combination of Python, Matrix Laboratory (MATLAB), and Interactive Data Language (IDL) [151]. CCDC/COLD is also available in MATLAB [339] and Python [119]; while it has been used less often that LandTrendr or BFAST/BFAST Monitor, it is a newer method and the provided MATLAB implementation is very specific to Landsat data. Many of the phenological change detection studies use TIMESAT for smoothing and extracting phenological parameters, which is freely available online (after completing a short form) with excellent documentation [147].

The utility of easy availability in a change detection method is therefore that the method will be used more frequently. Given the difficulties in comparing between methods, re-use of methods for multiple analyses in different regions of the world and to different land use types is desirable because it allows for more meta-analysis between studies. While diversity in change detection is desirable, having well established, standardised and widely used methods for particular applications would make results far more comparable, and also more easily replicable by other groups. This could be achieved through more open distribution of both implementations (ideally in popular languages for statistical analysis such as MATLAB, Python, and R) and source code.

2.2 Mangroves as a case study

To properly investigate the utility of a modelling approach requires a real world use case. This section describes the distribution and attributes of mangrove forests, their global importance, and current threats to mangrove ecosystems. The section describes why there is a need for mangrove monitoring which time series modelling approaches are well placed to meet.

2.2.1 Distribution and species

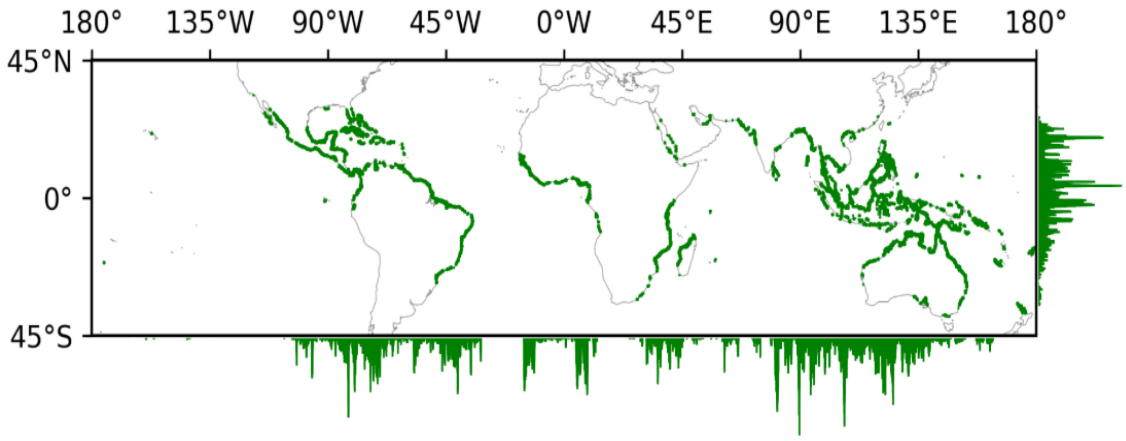


Figure 2.4: Global distribution of mangroves in 2010 according to the Global Mangrove Watch [39].

Mangrove forests grow in the tropical and sub-tropical regions of over 120 countries and territories worldwide [160] covering a total area of nearly 140,000 km^2 [39, 94]. As salt-tolerant evergreen trees, mangrove forests occupy a narrow ecological niche, colonising coastlines between around 32°N and 45°S (Figure 2.4). These intertidal regions represent a harsh growing environment, characterised by constant inundation with seawater alongside strong winds, high temperatures, and anoxic soils [160]. Mangroves have developed a number of morhpological, biological, physiological, and ecological adaptions to these conditions, including pneumatophoric (aerial) roots (Figure 2.5) and the ability to excrete salt through the leaves [160]. Many mangrove

species are also viviparous, with seeds that germinate and grow into seedlings while still attached to the parent plant before being dropped and dispersed by the tide [9].

Figure 2.5: Mangroves growing in the Dominican Republic. Photo by Anton Bielousov, distributed under a CC BY-SA 3.0 license [52].

Estimates place the number of mangrove species at around 50-70 [9, 160], with the highest concentration being in Asia which accounts for nearly 40% of the world's mangrove forests [39, 160]. The height of mangrove trees is highly variable depending on conditions, from 1 m high mangrove shrubs to 40 m high canopies, with biomass generally decreasing with increasing latitude [9, 160].

2.2.2 Significance of mangrove ecosystems

Globally mangrove forests provide a range of ecosystem goods and services and are of high biological, economic, and ecological importance [94, 160]. Mangrove forests represent some of the most biodiverse regions of the planet [88, 103], providing habitat, nurseries, and breeding sites for a range of marine and terrestrial fauna including crabs, prawns, fish, birds, reptiles, and mammals [9, 160]. Mangroves are also a source of food, timber, and traditional medicines and play a role in shoreline stabilisation and water filtration via their complex root systems [23, 172, 188]. In addition,

mangroves help to absorb the impact of extreme weather events such as cyclones and tsunamis, breaking up the impact of waves and wind [188].

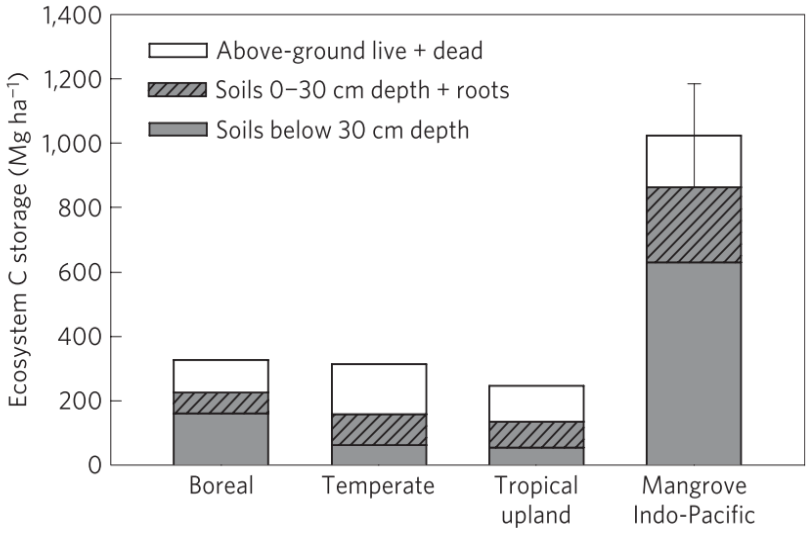


Figure 2.6: Carbon capture of mangrove forests compared with other forest types. From [68].

Mangrove forests are also important carbon sinks [158]. While above ground C storage is similar in mangrove forests to other forest types, the thick, organic-rich mud which characterises the ground layer of mangrove forests accounts for up to 98% of mangrove C storage [68]. This makes mangrove forests some of the most carbon-dense tropical forests (Figure 2.6), containing on average 1023 Mg of C per hectare and representing as much as 10% of carbon emissions from deforestation while only accounting for 0.7% of tropical forest globally [68].

2.2.3 Threats

There is a large body of research that suggests mangrove forests are under threat from a variety of local and global sources [74, 94, 160, 178, 306]. A study by Valiela *et al.* [293] estimated that around 35% of mangrove forest worldwide has been lost since the 1980's, with other studies suggesting an ongoing loss of 1-2% per year [74, 236], though this is difficult to quantify accurately given the scale and diversity of

mangrove forests. Richards and Freiss *et al.* estimated that between 2000 and 2012 more than 100,000 ha of mangrove forest was lost in Southeast Asia alone.

The most significant threat on local scales is loss of mangroves due to expansion of aquaculture practices [160]. This typically means the felling of mangrove stands to create ponds for raising fish or shrimp [254] and has been estimated to account for around 38% of mangrove forest loss worldwide [236]. In addition, Polidoro *et al.* found that over-exploitation for fuel and timber has resulted in the degradation of a quarter of the world's mangrove forest, with 16% mangrove species now being at risk of extinction [236]. Mangroves continue to be threatened by encroachment of human activity, including increasing human populations, industrialisation, urbanisation, pollution, and exploitation of natural resources such as oil [160, 201, 236, 253].

On a broader scale, mangroves are also under threat from the effects of anthropogenic climate change. Mangrove forests are generally resilient to short-term environmental change such as that caused by tidal inundation, seasonal rainfall, and storms [178, 212]. However, this resilience is at risk of being eroded. Rising sea levels mean that mangroves are inundated more frequently with saltwater and at greater levels of salinity, causing forests to retreat from the shoreline as their tolerance is exceeded [178, 306]. Climate change is also likely to lead to increasing numbers of extreme weather events such as tropical cyclones [126], which can have direct and devastating impacts on mangrove communities. For example, the impact of Cyclone Sidr on the Sundarbans mangrove forest in 2007 was estimated to have caused damaged to more than 20% of the forest [6] and super typhoon Haiyan caused widespread destruction to mangroves in the Philippines when it made landfall in 2013 [299].

It should be noted that there are also potential benefits to mangrove communities as a result of climate change. For example, there is evidence that rising sea levels can lead to inland migration of mangroves. Krauss *et al.* found that mangrove numbers

in southwestern Florida, USA increased by approximately 35% between 1927 and 2005, partly as a result of increased water depth further inland allowing mangroves to displace salt marsh [156]. There is also evidence that climate change is affecting mangrove range, allowing for poleward expansion [260]. However, a 2017 review by Feller *et al.* found that climate change played a complex role in mangrove expansion and retreat, with many areas experiencing simultaneous loss and gain [74].

2.2.4 Change in mangrove ecosystems

Figure 2.7: Mangrove seedlings colonizing mud cracks on the coast of French Guiana. This region experiences rapid fluctuations in mangrove extent, due to the constant formation and dissipation of mud banks caused by outflow from the Amazon Basin. A more established mangrove stand is visible in the background. Adapted from Proisy *et al.* [241].

Mangrove ecosystems are highly dynamic and may be affected by multiple change drivers, both natural and anthropogenic, and often occurring at different spatial and temporal scales. While monitoring extent (i.e. area covered by mangroves) is important, it is beneficial to also consider subtler changes in mangrove forest condition,

such as changes in height, photosynthetic activity, canopy cover, and species composition. These changes do not indicate a change in land cover class, but can be used to monitor longer term trends in mangrove forest health which may lead to forest loss or mangrove recovery and establishment. For example, monitoring the establishment and growth of mangrove saplings is key to determining the timeline of colonization. As pioneer species, mangroves often expand outwards by spreading onto newly formed land. Figure 2.7 shows an example of rapid mangrove colonization in French Guiana. In this region, large-scale deposition of sediment from the Amazon Basin leads to the formation of large mud banks up to 60 km long and 30 km wide [87]. The existence of these mud banks leads to creation of intertidal mud flats which experience rapid mangrove development followed by rapid erosion as the bank itself degrades [10]. These rapid changes make the coastline of French Guiana one of the most dynamic mangrove coastlines in the world.

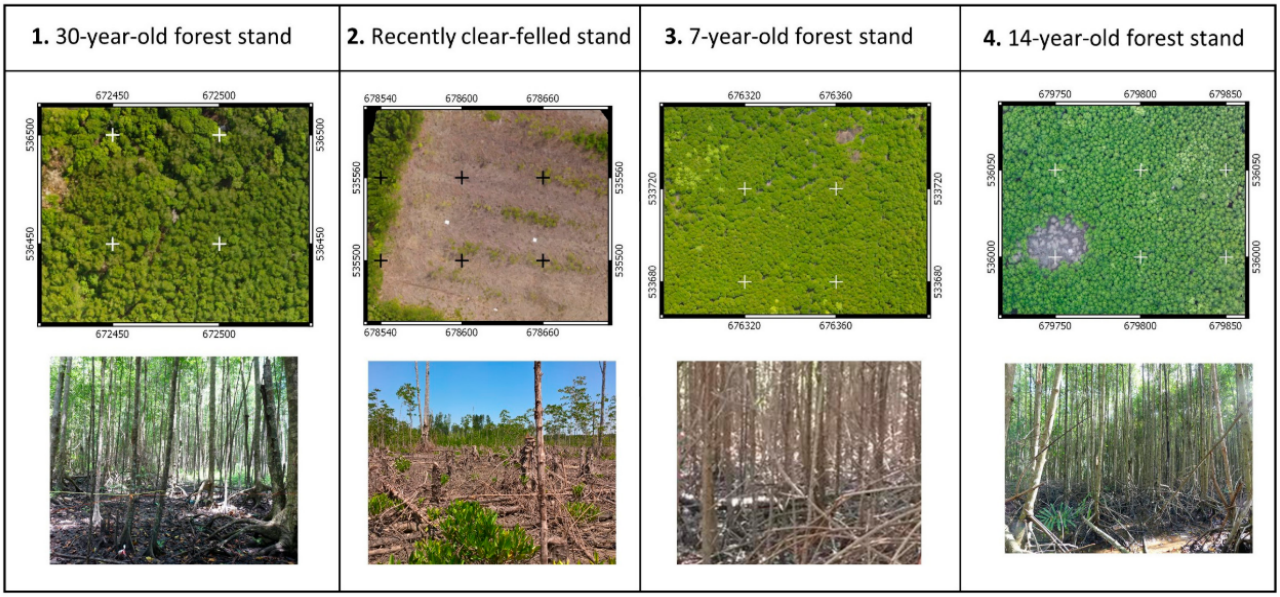


Figure 2.8: An example of forest clearing and regeneration in the Matang Mangrove Forest Reserve, Peninsular Malaysia. The mature forest stand (1) is felled as part of local forest management practices, leaving an area of exposed soil and debris (2). Within seven years, a dense canopy has re-formed (3) and remains robust to localized disturbances and partial thinning (4). Adapted from Otero *et al.* [217].

In some regions, mangrove condition change is directly driven by anthropogenic

factors. For example, the Matang Mangrove Forest Reserve is frequently used as a case study (e.g. [100, 127, 213]) because local silviculture practices mean that there are records of mangrove clear felling and regeneration (Figure 2.8). Current evidence is that this practice is sustainable and is not leading to any net loss of mangrove forest [1, 100]. However, the availability of data in the region makes it an excellent case study for determining the speed of regrowth and canopy closure, and robustness of mangrove stands to localized change. Much more common is the practice of felling mangrove trees to make room for aquaculture ponds (Figure 2.9), typically used for farming of fish, shrimp, and crabs. In these cases the loss of mangroves is intended to be permanent, however there is evidence that mangroves can re-colonize abandoned ponds [346]. Considering that the average lifespan of aquaculture ponds is around 10 to 13 years [14], re-establishment efforts will be vital to restoring mangrove populations in regions where mangrove forests have been decimated by aquaculture.

Figure 2.9: Clear felling of mangroves to create aquaculture ponds in Southeast Asia [233].

A direct example of mangrove condition change due to human activities is the

impact of oil spills. Crude oil has dramatic and devastating impacts on vegetation. Oil and oil by-products contaminate soil and coat mangrove roots and leaves hindering the exchange of gases, water, and nutrients [215]. The impact of spills can be devastating, causing death in even mature mangrove trees within as little as six months [215]. Unlike when mangroves are cleared for aquaculture, in this case mangrove loss is purely a by-product of accidental spills with no direct human benefit. The Niger Delta presents a specific case where monitoring of mangrove condition change from oil spills would be highly beneficial. The area contains Africa's largest contiguous mangrove forest (the third largest globally) and is also one of the world's top greenhouse gas emitters [215]. It is estimated that between 9 and 13 million barrels of oil have been spilled in the Niger Delta since 1958 [22], causing widespread damage to mangroves (Figure 2.10). However, the region remains under-studied in terms of mangrove forest change, and in particular there are no current estimates of either mangrove forest loss or the cumulative effects of oil spills on mangrove health. It is likely that low-level contamination has effects on mangrove health even in areas that avoid direct spills. However, such long-term effects on mangrove condition are impossible to monitor on the ground, and observations from satellites are hindered by cloud cover.

As mentioned in Section 2.2.3, cyclones can cause major damage to mangroves, and cyclone events are likely to increase in future as a result of climate change [126]. Direct impact by cyclone winds can flatten large areas of mangrove forest, and less severe impacts can strip trees of their leaves or cause inundation by floodwater leading to leaf browning as a result of salinity changes (Figures 2.11 and 2.12). Mangroves may recover from these more minor effects, but it has been suggested that the cumulative impact of cyclones on areas such as the Sundarbans is impacting mangrove resilience and reducing the protective capacity of the forest [272]. A 2020 study by Mandal *et al.* estimated that at least 21 cyclones have directly impacted the Sundarbans since the late 1980's, with some cyclones affecting nearly a quarter (24.1%) of the total

Figure 2.10: Mangrove degradation caused by oil spills in the Niger Delta, Western Africa [215]. As of 2020 there were reported to be over 900 active oil wells in the region [215].

forest area. However, so far the majority of studies looking at the Sundarbans have focused on extent rather than attempting to monitor mangrove health directly (e.g. [60, 98, 244]) though Ghosh *et al.* studied long-term changes in species composition [91].

Further evidence for the cumulative effects of environmental stress on mangrove health is the large-scale dieback event which occurred in the Gulf of Carpentaria in late 2015. Despite being relatively undisturbed for the last 200 years, over 7000 ha of mangroves along 1000 km of coastline suddenly died with no obvious cause [69, 112]

Figure 2.11: Examples of damage to mangroves caused by Cyclone Sidr, adapted from Akhter *et al.* [6]. Cyclone Sidr made landfall in Bangladesh on the 15th of November 2007, directly impacting the Sundarbans mangrove forest and causing severe damage to vegetation, wildlife, and infrastructure. The cyclone was estimated to have damaged 22% of the total forest area [6].

Figure 2.12: Flooding in the Sundarbans caused by the impact of Cyclone Aila in 2009. There is concern that repeated cyclone impacts could reduce the protective capacity of mangrove forests such as the Sundarbans. Image by Arun Sarkar [272].

(Figure 2.13). Subsequent reports suggested that there were signs of poor health in the region's mangroves throughout the previous year, and the dieback event was likely due to a combination of environmental stressors including high temperatures and low sea levels [69, 112]. Lymburner *et al.* noted that despite the dieback event, total mangrove extent in the area had still increased overall between 1992 and 2016 [182]. However, mangrove recovery in the area remains at risk, with small re-establishing mangrove stands being particularly vulnerable to cyclone impacts [69]. Longer-term

monitoring is needed to both establish a baseline of mangrove extent in the area and to assess recovery over the next 5-10 years. With improved assessments of mangrove health, such events could be anticipated in future if unusual fluctuations or rapid declines in mangrove health can be observed.

Figure 2.13: Damage to mangroves as a result of a dieback event in the Gulf of Carpentaria, Queensland, Australia in 2015. A 2017 report concluded that the damage was likely caused by cumulative stress due to unusually low sea levels and high land surface temperatures [112]. Image by Norman Duke [347].

Large scale studies have typically focused on mangrove extent change, as condition change is difficult to monitor over large areas. The Global Mangrove Watch, for example, aims to create accurate global mangrove maps [39] from which yearly changes in extent can be determined. However, metrics such as the NDVI which act as a proxy for photosynthetic activity can be used to monitor trends in mangrove forest health. This has been done at local scales. Long *et al.* used NDVI to monitor damage and recovery in mangroves in the Philipines following Super Typhoon Haiyan, by using NDVI to classify mangroves into damage level categories [177]. A similar approach was taken in an earlier study by Datta *et al.* in the Sundarbans [60], and Duke *et al.* used differences in NDVI between years to identify areas of change in the Gulf of Carpentaria [69]. Cornforth *et al.* used SAR to observe changes in mangrove forest

structure [55]. However, these studies still focus on specific thresholds or classification and differencing and are limited to inter-year comparisons. The approaches discussed in Section 2.1.2 such as EWMACD/Edyn and COLD have the potential to provide detailed, long-term information on changes in both mangrove extent and mangrove condition while also taking account of ephemeral change. By taking a model fitting approach that makes use of all available observations, trend information can be extracted independently of seasonality or tidal fluctuations, providing greater insight into both the timing and causes of mangrove forest change.

2.2.5 Remote sensing for mangrove monitoring

The global significance of mangrove ecosystems alongside the complex local and global interactions between mangroves, human activities, and climate means there is a need for accurate mangrove monitoring over long time scales. However, mangrove forests are often extensive, highly dynamic, and difficult to access, meaning that field studies are expensive, time consuming, and impractical [91, 97]. Obtaining a near-simultaneous assessment of an entire mangrove forest on the ground is impossible, yet even more vital in the case of mangrove forests where there is constant erosion and deposition of sediment causing mangrove expansion and retreat, sometimes on very rapid timescales [89]. Remote sensing is the obvious solution to this problem and the utility of remote sensing for mangrove monitoring was realised as early as 1968 [51]. Currently, satellites such as the Landsat and Sentinel missions capture medium-scale spatial resolution data over large tracts of land on daily to weekly timescales, allowing for ongoing monitoring of large tracts of land in addition to examination of past data to determine ongoing trends. The Landsat archive, going back to the 1970's, is a particularly useful resource for studying long-term change.

Figure 2.14 shows a steady year-on-year increase in the number of studies using remote sensing for mangrove monitoring over the last 20 years. Over the last two

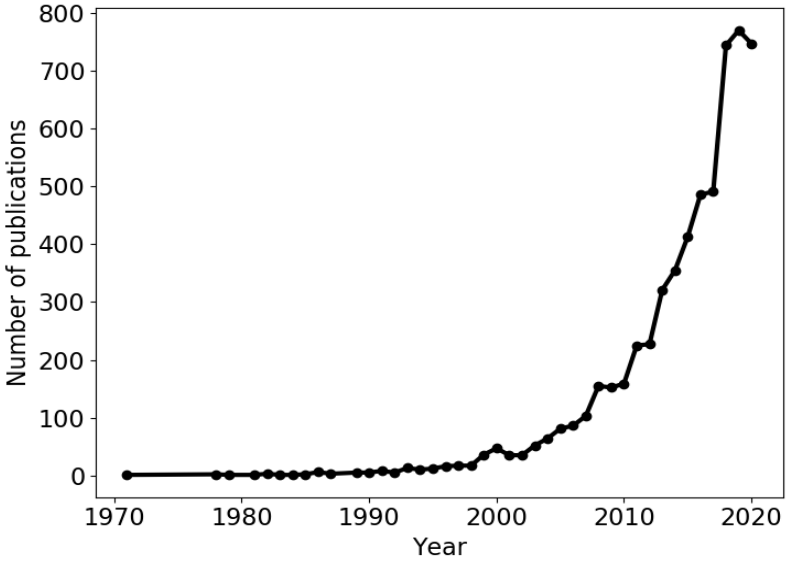


Figure 2.14: Plot showing the number of articles published per-yer containing the words "mangroves" and either "satellite" or "remote sensing", in the subject area of Earth sciences. Data retrieved from Scopus.

decades, satellite data has been used to monitor mangrove forest extent (e.g. [98, 127, 182]), structure (e.g. [131, 180, 284]), biomass (e.g. [92, 240, 315]), and species composition (e.g. [91, 99, 127]) using both optical (e.g. [2, 13, 225, 256]) and radar (e.g. [55, 240, 286]) data. Reviews of remote sensing for mangrove monitoring were published in 1998 [106], 2011 [116, 160], 2016 [95], 2017 [324], and 2019 [230, 305], leaving no doubt that this is an intensely active area of research. However, challenges remain, particularly when considering global-scale, long-term studies of mangrove extent and condition. Very few studies have taken advantage of the full Landsat archive for long-term monitoring; Younes Cárdenas *et al.* looked at 50 studies and found that the majority used five images or less [324] and that more than 50% of the studies reviewed focused on just five countries (Australia, Malaysia, China, Madagascar, and Mexico). Some areas, notably West Africa and parts of South America, remain extremely under-studied [324].

The spatially and temporally fragmented nature of mangrove research means it is difficult to make comparisons between different regions or to gain global insight into

the condition of mangrove forests. However, while the need for a global mangrove product has long been identified [117, 305], such products remain rare and either only provide extent for single year [94] or provide year-to-year comparisons based on composite imagery, often with large gaps in the change record [39, 286]. The Global Mangrove Watch (GMW) represents the current state of the art in global mangrove monitoring. Utilising mainly Landsat and ALOS-PALSAR data, the GMW has produced a baseline mangrove extent map for 2010 in addition to estimates of mangrove change from the mid-1990's onwards. While extent estimated by the GMW is generally considered reliable, the methodology used is time consuming and involves extensive manual quality assurance [39]. To address the issue of incomplete imagery the GMW used an image compositing approach.

Another limitation of remote sensing for mangrove monitoring lies with the data record. While the Landsat archive contains decades of data, for many tropical regions cloud cover renders the majority of that data unusable [160]. Even where cloud cover is not an issue, seasonal variation requires the use of comparison images from the same time of year. For this reason studies tend to focus on comparisons between specific, cloud-free images (e.g. [43, 91, 127, 159, 207]) or use compositing approaches (e.g. [213, 217, 256]). Change estimates based on such imagery are limited in the case of mangrove ecosystems which may change rapidly due to tidal and seasonal variations. However, in recent years the field of land cover monitoring has seen seen a move away from year-to-year comparisons and towards time series approaches which make use of all available observations [338]. Methods such as Breaks for Additive and Seasonal Trend (BFAST) [296], Exponentially Weighted Moving Average Change Detection (EWMACD) [35], and Continuous Change Detection and Classification (CCDC) [340] provide a potential solution to the problem of global, long-term monitoring of mangrove extent. As discussed in Chapter One, such approaches use temporal models to capture the seasonal dynamics of land cover types in addition

to longer term greening and browning trends. These methods have particular utility for mangrove monitoring, because they reduce the reliance on single observations in time, instead making use of every available observation in the time series. This allows for satellite images which are nearly completely obscured by cloud to be included in analysis on the basis that there is at least a small amount of useful data present. In addition, such modelling approaches have the potential to capture the full variability of the underlying land cover, allowing for more accurate monitoring of change.

Despite the advantages of time series modelling, uptake of these methods for tropical forest monitoring has been slow and at the time of writing no studies exist which make use of time series modelling for mangrove monitoring. One reason for this is that until very recently the computing resources needed to apply such methods were simply not available to researchers. Time series models are generally applied on a per-pixel basis and involve downloading and processing hundreds, if not thousands, of satellite images. This requires large amounts of both storage and processing ability. Platforms such as GEE and Amazon Web Services (AWS) are reducing these barriers [95]. GEE in particular is designed for large scale spatial and temporal satellite image analysis and many time series modelling approaches - such as CCDC, LandTrendr, and EWMACD - have implementations in GEE. A 2019 study by Mondal *et al.* used GEE to test two methodologies for mapping mangroves in Senegal [203]. Another 2019 study used GEE to map mangroves in China using Sentinel-2 imagery [168] and a 2020 study used GEE to study mangrove damage from cyclones in the Sundarbans region [186]. However, none of these studies utilise a time series modelling approach, though a 2019 study by Tang *et al.* did investigate the use a CCDC-style algorithm for tropical forest monitoring [283].

2.2.6 Summary

There remains a need for a globally consistent methodology for monitoring mangrove extent and condition which spans the full range of the available data archive; however, any globally applied methodology must be able to deal with the issue of low data coverage which exists for many regions occupied by mangroves. Model-fitting approaches provide an obvious solution to this problem, as they can interpolate between observations and make use of the full data record. These methods can also be used to extrapolate information other than extent (e.g. trends in vegetation condition). However, based on a review of the literature, there has been very little research into the use of pixel-based time series modelling approaches for mangrove monitoring. In addition, as access to both online and local high performance computing capabilities continue to improve, use of larger satellite image data sets will become easier and more frequent. Mangrove monitoring is therefore used here as a use-case to test the efficacy, practicality, and accuracy of temporal modelling for land cover monitoring. For simplicity and to reduce the computational burden, this thesis will focus on a per-pixel modelling approach. Since the pixel is the smallest unit of spectral signal this offers substantial benefits over object-based approaches in terms of spatial detail, and reduces any complexity involved with incorporating change information from the surrounding area. Per-pixel approaches are also advantageous in that they are very easy to parallelize on HPC systems, as each pixel can be considered as an independent unit. This simplifies computation when applying an algorithm over a large spatiotemporal dataset.

2.3 Research questions

This thesis will aim to answer a set of research questions arising from the literature review.

- Can simulated remote sensing time series provide insights into the efficacy of change detection methods by providing an objective measure of change?
- What are the strengths and limitations of the popular change detection methods known as BFAST, BFAST Monitor, CCDC, and EWMACD/Edyn?
- Based on their strengths and weaknesses are there specific and appropriate use cases for each method?
- Which of these method(s) are best suited to the use-case of mangrove monitoring, given the challenges involved?
- Can these methods be used to estimate trends in vegetation condition (e.g. degradation and recovery) in addition to changes in extent?
- What benefits are there to using a time series approach for mangrove monitoring compared to other approaches? Do these methods lead to more accurate estimates of mangrove extent and/or provide more information on mangrove change dynamics?
- Are cyclones in the Sundarbans having a cumulative long-term effect on mangrove ecosystem resilience?
- Are mangrove populations in the selected study areas increasing or decreasing? What does this tell us about mangrove population change globally?
- Is the application of a per-pixel time series modelling method to mangrove monitoring feasible at the global scale?

Chapter 3

An Evaluation and Comparison of Four Dense Time Series Change Detection Methods Using Simulated Data

Before examining the specific use-case of mangrove forest change, Chapter 3 will investigate the general suitability of time series modelling methods for change detection in remote sensing time series. Simulated data sets will be used to test the performance of four methods (described in detail in Chapter 2) on a range of land cover change scenarios. From this, general conclusions will be drawn on the strengths and weaknesses of each method. The ability of these methods to distinguish change in vegetative land cover types will be discussed, before identifying the method most appropriate for mangrove forest monitoring.

Land use type contributes to anthropogenic climate change by impacting photosynthetic activity, transpiration, and albedo. It has been suggested that agriculture, forestry and other land use change could account for 21% of anthropogenic greenhouse gas emissions [290]. Van der Werf *et al.* estimated that 6–17% of anthropogenic CO_2 emissions could result from deforestation alone [310]. As such, the ability to accurately monitor land use and land cover change can be pivotal in understanding and mitigating the effects of climate change.

The launch of the Landsat 8 mission in 2013 [258] and the Sentinel-2 missions in 2015 and 2017 resulted in an increase in available optical satellite data with 5–16 day temporal resolution. Such temporally dense time series provide the opportunity to capture the complex seasonal dynamics of many land cover types and to detect land cover change more rapidly than ever before. In addition, the opening of the Landsat archive in 2008 provided access to nearly 40 years' worth of free historical data [319]. Methods such as LandTrendr [150], Composite2Change [114], Vegetation Change Tracker [122], and ShapeSelectForest [202] have been developed to exploit the Landsat data archive to examine long-term vegetation trends. However, these methods focus on comparing yearly composite images. The focus of many land use change detection studies has now shifted towards detecting change on a per-acquisition rather than a yearly basis, with new methods being developed to exploit these temporally dense time series by using season-trend models to account for intra-year variability [338]. An early example of this is Harmonic Analysis of Time Series (HANTS), which uses an iterative season-trend modelling approach for time series smoothing and interpolation [255]. In addition, Saxena *et al.* [263] demonstrated that combining the output of several methods in an ensemble approach can produce a more accurate result. However, effectively selecting which methods to use or combine requires knowledge of each respective method's strengths and weaknesses.

Given that demand for dense time series monitoring is only likely to increase, emphasis must be placed on evaluating the temporal accuracy of land use monitoring methods. However, this is not a straightforward process. Specifically, it can be difficult to find appropriate ground truth datasets where the date of disturbance is precisely known, with many studies relying on labour-intensive human interpretation of data to produce a validation set. Tools such as TimeSync [49] are growing in popularity (e.g., [50, 70, 263]) and can aid accurate signal interpretation by allowing users to view and classify pixels within their spatiotemporal context, alongside

higher resolution data from Google Earth [49]. Despite such tools, the availability of reliable change validation datasets remain scarce and such datasets will always be prone to human error. Furthermore, change detection studies tend to be focused on particular types of changes, with an *a priori* understanding of break magnitudes or underlying trends. These limitation make it difficult to develop and evaluate universal approaches to change detection, or draw comparisons between different methods of change detection.

Given the difficulties in obtaining suitable "real-world" data to evaluate and compare change detection methods the use of simulated time series data offer a tractable solution. Simulations can be easily generated in large numbers, can contain fixed changes of known magnitude, can include multiple types of change, and can include known quantities of noise or missing data. Despite these advantages, few studies have used simulated data in remote sensing. Studies such as those by Verbesselt *et al.* [296] and Forkel *et al.* [75] have used simulations to evaluate new methods and compare between methods, respectively. However, in the case of Verbesselt textitet al. only one method was being evaluated, whereas Forkel *et al.* only focused on changes in trend. No studies exist which have aimed to comprehensively compare multiple change detection algorithms across a wide variety of change types.

This chapter compares four popular change detection methods: (1) Breaks for Additive and Seasonal Trend (BFAST); (2) BFAST Monitor; (3) Continuous Change Detection and Classification (CCDC); and (4) Exponentially Weighted Moving Average Change Detection (EWMACD). Comparisons are made using simulated NDVI data representing a range of change types and magnitudes. The effectiveness of each method was analysed in multiple areas including efficacy at detecting true changes, likelihood of detecting false changes, response to noise, response to missing data, and accuracy in determining the magnitude of a change.

3.1 Materials and Methods

3.1.1 Change Detection Methods

The aim of this study was to compare and evaluate a range of methods used for change detection analysis of temporally dense satellite image time series. To achieve this, four approaches were used: BFAST, BFAST Monitor, CCDC, and EWMACD. These four approaches all use a season-trend decomposition model to take account of both inter- and intra-year variation in a time series. Changes are found by determining where in the time series a model breaks down and no longer adequately fits the data, indicating a change in land cover. A new model can then be fitted to the next period in the time series.

The intention was to investigate the off-the-shelf performance of these methods, rather than tailoring them to any particular scenario, to obtain a broad assessment of performance. Each method has its own user-definable parameters and where possible either default values or values which facilitated comparability across methods were used. As a result, performance is likely to be poorer in some cases than could be achieved with more parameter tuning. Each method along with parameters used is outlined in detail below. The scripts used to run each method on the simulated datasets are available at [17].

3.1.1.1 BFAST

The BFAST R package was used in this study [298]. BFAST is a widely used method for detecting trend and seasonal breaks in time series. It has mainly been applied to monitoring forest disturbance (e.g., [70, 109, 264]) but has also been applied to more general land cover monitoring scenarios (e.g., [44, 235, 307]). BFAST uses an iterative process to find both trend and seasonal changes across a whole time series [296]. It should be noted that a trend change here refers to an abrupt change in the trend of the time series, rather than a gradual slope. First, an Ordinary Least Squares

Moving Sum (OLS-MOSUM) test is used to determine if any breakpoints are present in the time series. If the OLS-MOSUM test indicates significant ($p < 0.05$) change, the number and location of breakpoints is estimated separately for the seasonal and trend components using OLS fitting. The BFAST package automatically fits a third-order harmonic model. The result is a set of piecewise season-trend models which minimise error across the whole time series. The difference between the intercept and slope terms of consecutive models is used to calculate change magnitude between breakpoints [296].

BFAST requires two user-defined parameters: (1) the minimum distance between breaks; and (2) the maximum number of iterations. Saxena *et al.* [263] suggested that the number of breakpoints is the most influential parameter—if the number of breaks exceeds the number of breaks defined by the user, then it will only find the strongest. The minimum distance between breaks was set to two years (46 observations), which is in line with the guidelines given by Verbesselt *et al.* [296] and matches the two-training period used for the other methods. BFAST requires that time series have no gaps so linear interpolation was used for simulations with missing data.

Initially, BFAST was allowed to run for up to 50 iterations, but testing showed that in most cases convergence was achieved after five iterations. In other cases, convergence was still not achieved after 50 iterations. Given that runtime increases significantly with the number of iterations, computational efficiency was balanced with an adequate number of outputs that achieved convergence by setting the maximum number of iterations to five.

3.1.1.2 BFAST Monitor

BFAST Monitor was developed as a near-real time alternative to BFAST [297]. Similar to BFAST, it has mainly been applied to forest monitoring [65, 206, 267]. It is based on the premise that change can be identified by looking for deviation of new observations from a stable history period. Unlike BFAST, BFAST Monitor does

not attempt to separate seasonal and trend changes. The season-trend model given by Equation (3.1) is fitted to the stable history period using OLS. Here, y_t represents the data at time t, α_1 is the intercept, $\alpha_2 t$ is the slope, k is the number of harmonic terms, $\gamma_1,\ldots,\gamma_k$ represent the amplitudes, $\delta_1,\ldots,\delta_k$ represent the phases, f represents the number of observations per year, and ε_t is the error. When new observations are available, residual values are calculated using the fitted model and Moving Sums (MOSUMs) of the residuals are used to look for instability which would indicate structural change [297]. This allows BFAST Monitor to flag a change within a single observation.

$$y_t = \alpha_1 + \alpha_2 t + \sum_{j=1}^{k} \gamma_j sin\left(\frac{2\pi jt}{f} + \delta_j\right) + \varepsilon_t \tag{3.1}$$

BFAST Monitor was run using the R package [298]. Given that all simulations were designed with a break after five years of stability, a stable history period of two years (46 observations) was used. While this could have been longer, allowing three years of data between the end of the history period and the true date of change allowed for assessment of how likely the methods were to find false breaks. A second-order harmonic model was chosen for BFAST Monitor because that is the maximum complexity of the simulations used. Unlike BFAST, BFAST Monitor can be used on datasets with missing values. BFAST Monitor uses the difference in medians between the history period and monitoring period to estimate break magnitude [297].

The R implementation of BFAST Monitor does not allow for continuous monitoring. Therefore, a process was implemented whereby, after a break is detected, if at least 46 more non-missing values are available, BFAST Monitor is re-run with the new history period until either another change is found or the end of the time series is reached.

3.1.1.3 CCDC

CCDC focuses on changes in land cover class [340]. However, the classification component was not used here because the simulated data were not designed to relate directly to specific land cover types. Similar to BFAST Monitor, CCDC aims to detect changes in near-real time. The model used by CCDC is very similar to the season-trend model used by BFAST Monitor, except that CCDC uses an adaptive process to minimise model overfitting while also robustly capturing the seasonal cycle [342]. Rather than using a fixed number of coefficients, CCDC fits a second-, third-, or fourth-order harmonic model depending on how many observations are available in the training dataset [342]. To avoid overfitting of higher-order models, Lasso regression is used instead of OLS to fit the season-trend model to the history period. Lasso regression minimises overfitting by limiting the total absolute value of the coefficients [342]. As a result, some coefficients will be forced to zero and will have no influence on the model.

The version of CCDC used here is based on that of Zhu *et al.* [342], where six new observations are needed to reliably flag a change from the stable history period. Change is identified using the Root Mean Square Error (RMSE) of the fitted historical model and the residuals of the incoming data. If the new residuals deviate from the fitted model six times in a row, the date of change is identified as the date of the first deviation and change magnitude is the residual value for that date. Once a change is identified a sliding window approach is used to determine the next stable period [340]. At the time of conducting the study, there was no freely available implementation of CCDC suitable for use with simulated data so a suitable implementation was written in the Python programming language.

All of the tested change detection methods rely to some extent on parameter tuning to achieve the best results. Due to the use of Lasso fitting, CCDC is less reliant on the user to choose the number of harmonics or the length of the history

period. However, Lasso regression has the potential to provide much finer grained control over model fitting by setting the parameter λ, which controls the degree to which Lasso penalises the coefficients. While a fixed value of λ can be used [66, 344], a substantially better result could potentially be achieved using a cross-validated approach. Cross-validation can be used to find the optimal value for λ by fitting multiple models with different values and comparing them. These two approaches are referred to as CCDC and CCDC with Cross Validation (CV). For the fixed approach, a value of $\lambda = 0.01$ was chosen based on small scale testing. Other studies have reported values of 20 [66, 344]. However, in these cases, the models were being fitted to surface reflectance or similar products, rather than NDVI.

3.1.1.4 EWMACD

EWMACD specialises in subtle changes, such as partial changes within pixels [35]. Unlike the other three methods, EWMACD also detects condition (increasing/decreasing trend) changes because it only fits a seasonal model without a trend term. EWMACD uses a specific type of statistical control chart, the EWMA chart, to rapidly find changes in time series. Statistical control charts were developed as a form of quality control in manufacturing and use *control limits* to establish when a time series deviates from a stable state. The Moving Sum (MOSUM) and Cumulative Sum (CUSUM) charts used by BFAST and BFAST Monitor are other examples of statistical control chart.

EWMACD calculates the residuals for a given training period based on a seasonal model fitted with OLS. To match BFAST Monitor, a second-order seasonal model and a two-year history period were used. This produces a set of normally distributed, independent observations suitable for use with an EWMA chart. To produce the actual EWMA values, the residual for each time point is adjusted to be a weighted sum of all previous values where the degree of weighting is specified by a parameter $0 < \lambda \leq 1$ [35]. The closer the value of λ is to one, the less weight is given to historical

data. Following Brooks *et al.* [35], the default value of $\lambda = 0.3$ was used. Upper and lower control limits are then calculated based on the mean and standard deviation of the residuals and the value of λ. When new observations are available, they are added to the chart and if their value exceeds the upper or lower control limit for a specified number of times then the change is said to be persistent and is flagged. A value of six observations required to flag persistent change was chosen to match CCDC. Whilst EWMACD produces values for break magnitude, these are relative and could not easily be compared to the other methods and therefore the residual was used as with CCDC.

The freely available version of EWMACD was used in this study [36]. This version does not allow for continuous monitoring and therefore a later implementation of EWMACD called dynamic EWMACD (Edyn) [37] was used to add this as a feature. Edyn uses a vertex approach to determine where a time series re-stabilises after a break by finding the point of greatest deviation between the date of change and the most recent observation [37]. The algorithm is then re-run from the date of stabilisation. If any deviation is flagged in the new training period, a sliding window approach is used where one observation is removed from the front of the time series and one added to the end until a new two-year period with no flagged changes is found. A value can be provided to EWMACD to screen out erroneously low values (i.e., values below zero NDVI for vegetated pixels), but since our simulations include negative trends this was set to -1 to keep all observations.

3.1.2 Simulating Seasonal Time Series

The method used to generate the simulated NDVI time series was based on that described by Verbesselt *et al.* [296, 297]. NDVI was chosen because most of the methods are designed to work on a single band or index. Furthermore, NDVI is a well-recognised and widely used metric for examining trends in vegetated areas. The

method involves using a double Gaussian function to simulate an NDVI signal over time, where $t = 1,\ldots,t = n$ for a time series with n observations per year. For each year in the time series, the NDVI value at time t is defined by the amplitude of the seasonal curve (a) (i.e., the peak NDVI value), the base or lowest winter value, the location in time of the maximum value for each year b, the width of left left hand side of the curve (c_1), and the width of right hand side of curve (c_2) (Equation (3.2)). Based on work by Verbesselt *et al.* [296], a value of $b = 12$ was used to simulate 10-year time series at a 16-day temporal resolution giving approximately 23 observations per year and centering the curve around the middle of the year. The trend component is a small NDVI value which is added or subtracted cumulatively from each value, to create an upward or downward trajectory. The noise component was added randomly to create more realistic variation in the time series, as explained in Section 3.1.3.

The amplitude and width of the generated seasonal curve can therefore be altered by using the parameters a, c_1, and c_2. Increasing the value of c_1 results in a corresponding increase in Start of Season (SOS), as shown in Figure 3.1. The method described by White *et al.* [314] was used to calculate number of days by which the start of season had moved forward for the corresponding change in c_1 (Table 3.1).

$$f(t;a,b,c_1,c_2) = a \times \left\{ \begin{array}{ll} base + exp[-(t-b)^2/c_1] + trend + noise, & \text{if } t > b \\ base + exp[-(t-b)^2/c_2] + trend + noise, & \text{if } t < b \end{array} \right\} \tag{3.2}$$

Table 3.1: Table showing corresponding change in SOS for a given change in the c_1 parameter, based on [314].

Δc_1	ΔSOS
5	−13
10	−22
15	−30
20	−37
25	−43
30	−49

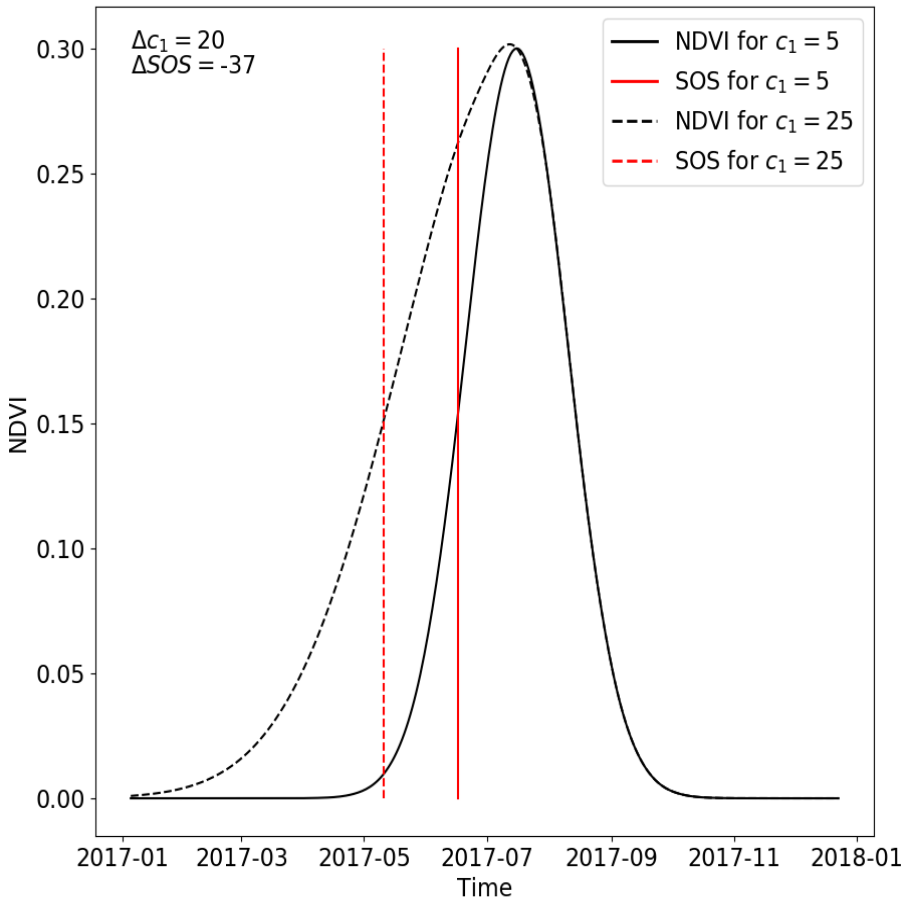


Figure 3.1: Plot demonstrating how a change in the c_1 parameter of Equation (3.2) results in a corresponding change in SOS. Here, the SOS has been moved forward by 37 days by changing the c_1 parameter from 5 to 25.

3.1.3 Noise

The presence of noise is inevitable in satellite image time series of optical data. Atmospheric and sensor effects can lead to random variation, which is difficult to screen

out. Robustness to noise is therefore important when considering which change detection method to use. Noise was added to the NDVI time series by randomly drawing a value from a normal distribution with a mean of 0 and a standard deviation of 0, 0.01, 0.02, ..., 0.07, meaning that a simulated time series with a noise level of 0.02 will have a random number between -0.02 and 0.02 added to each individual NDVI value. Therefore, it should be noted that simulations with higher noise levels have noise added from a wider distribution, and therefore contain both a wider variance of noise and higher individual noise values on average. Fifty simulations were generated for each level of noise in order to avoid any bias caused by noise being unevenly distributed throughout the time series (e.g., higher levels of noise being concentrated at the start of the time series).

3.1.4 Missing Data

Satellite image time series are rarely complete. The presence of contaminants such as clouds, cloud shadows, and snow causes anomalous values which can be detected and removed to some degree, leaving gaps. As with noise, robustness to missing data is therefore a crucial component of evaluation when considering change detection methods, especially when applying change detection to parts of the world with persistent cloud or snow cover.

Data were removed from each simulated time series by first calculating the number of observations to drop based on the length of the time series and the percentage data missing. This was rounded up to the nearest integer. A random number generator was then used to select observations based on their index in the time series. If an index came up more than once, the duplicate was discarded. NDVI values for the randomly selected indices were then removed. Fifty simulations were generated for each level of missing data to avoid any bias caused by missing observations being unevenly distributed throughout the time series.

3.1.5 No Change Set

A set of simulations was generated where no change occurred. This was done to assess how likely the different methods were to detect a change where none existed. These simulations maintain a consistent seasonal cycle throughout (Figure 3.2). The no change set consists of 2400 simulations (50 replicates for each of eight levels of noise and for each of five levels of missing data) (Table 3.2).

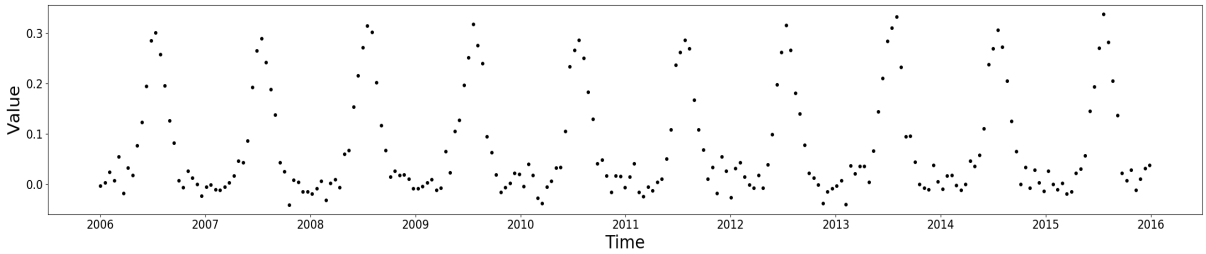


Figure 3.2: Simulated 10-year NDVI time series with a noise level of 0.02 and no missing data.

Table 3.2: All combinations generated for each level of noise and missing data. For the break/trend set, each abrupt change in NDVI is followed by either no trend or one of the six levels of trend present in the trend only set.

Simulation Type	Levels	No. Simulations
No change	—	2400
Trend only	0.002, 0.0015, 0.001, −0.001, −0.0015, −0.002	14,400
Break/trend	0.3, 0.2, 0.1, −0.1, −0.2, −0.3	100,800
Amplitude change	0.3, 0.2, 0.1, −0.1, −0.2, −0.3	14,400
LOS change	5, 10, 15, 20, 25, 30	14,400
NOS change	One to two, two to one	4800
Total	—	151,200

3.1.6 Trend Only Set

A set of simulations was generated which contains a constant negative or positive trend, but no other changes. This was done to assess how likely the different methods were to detect an abrupt change where none existed, if a constant trend was present in the time series. Long-term trends are often present for vegetative land cover types, for

example, due to land degradation [140] or the effects of global warming [75]. However, apart from EWMACD, all of the methods used in this study incorporate a trend term in the fitted model and are designed to flag only abrupt (step) changes or seasonal changes.

The trend only set consists of 14,400 simulations. Simulations were generated for six levels of trend (Table 3.2).

3.1.7 Seasonal Change Sets

A set of simulations was generated which contains a change in the shape of the seasonal curve. This was done to assess how well the different methods detect subtler changes in time series, in addition to abrupt/step changes. Given that all methods fit a seasonal component, it would be expected that fitted models would break down given a change in amplitude or Length of Season (LOS) because this would alter the fit of the model. However, BFAST is the only method which delineates seasonal changes from trend changes. The three seasonal change types are a change in the amplitude of the seasonal cycle (Figure 3.3), a change in the LOS, and a change in the number of seasons (i.e., from one peak per year to two) (Table 3.2). These simulation types were designed to imitate various changes in land productivity. For example, a change in seasonal amplitude or SOS (Start of Season) could indicate greater yield or an earlier planting, whereas a change in number of seasons simulates a change in number of yearly cropping cycles. The magnitude of the seasonal changes used was based on previous work by Verbesselt et al. [295–297].

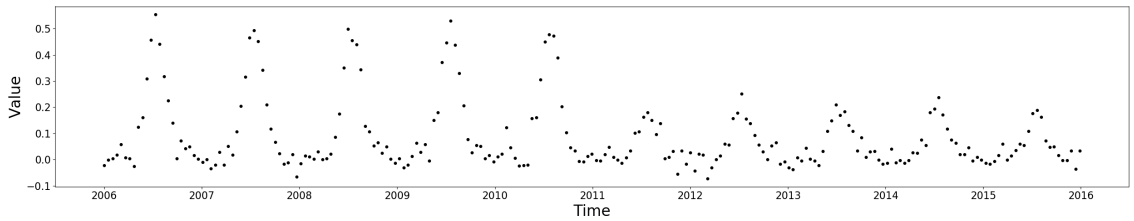


Figure 3.3: Simulated 10-year NDVI time series with a noise level of 0.03 and no missing data. A change in amplitude of -0.3 occurs halfway through the time series.

3.1.8 Break/Trend Set

A set of simulations was generated which contains different magnitudes of abrupt change followed by different levels of trend. This was done to simulate changes that occur from sudden events such as logging, fire, or flood, which may be followed by longer term recovery or degradation of vegetation. Different levels of trend were included because, while a trend should not be detected as a change in itself (except in the case of EWMACD), the presence of a trend after a break contributes to how easily the break is detected, especially if there is noise or missing data. For example, a low-magnitude positive abrupt change will be easier to detect if it is followed by a steep positive trend, because, even if the initial event is missed, the time series will continue to deviate significantly from the previous stable period. However, an abrupt drop followed by fast recovery, such as is shown in Figure 3.4, might be more difficult to detect because there is substantial overlap of the values from before and after the break.

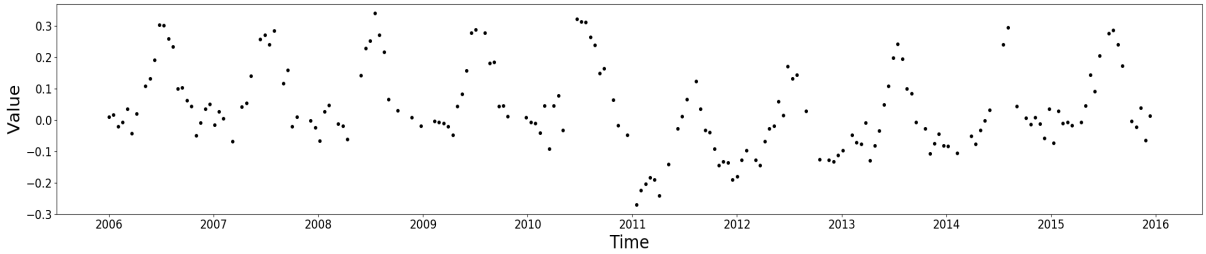


Figure 3.4: Simulated 10-year NDVI time series with a noise level of 0.04 and 20% missing data. A change in base NDVI of -0.2 occurs halfway through the time series, followed by a strong trend of 0.002.

3.1.9 Definition of Change

For the break/trend set, a correct change is defined as a change detected by the algorithm 96 days or less (equating to six observations, or roughly three months) after the date of the true break. Changes are always placed at the start of 2011 such that the earliest possible date the change could be detected given the data frequency

is the 15 January 2011. Given that the purpose of this study is to investigate the efficacy of these methods when applied to dense time series, detecting an abrupt change within a quarter of a year was considered a reasonable expectation. For the seasonal change sets, a correct change is defined as a change detected within one year (23 observations or 368 days), since changes in the shape or length of seasons are only of interest on a yearly basis.

3.1.10 Correlation Statistics

Correlation statistics were calculated using the non-parametric Spearman's rank measure. Spearman's ρ statistic provides an indication of the monotonic relationship between two variables (i.e., whether both variables increase or decrease together when ranked).

3.1.11 Computer Specifications and Timing

All steps including generation of simulations, production of results, and analysis were carried out on using a desktop computer with an Intel i7 CPU running at 4.20 GHz with 32 GB of RAM. Total runtime using a single process was recorded for each simulation set for each method. This total was then divided by the number of simulations in the set in order to obtain a mean runtime per simulation in seconds.

3.2 Results

3.2.1 Runtime

There was a lot of variation in how quickly the different methods processed the simulations. Table 3.3 shows that the sets with no changes generally took less time to process than those with changes, except for in the case of CCDC with CV, where they took longer. BFAST also took less time to process time series with NOS changes than those with no changes at all. There is a clear difference between CCDC and CCDC

with CV, with the latter taking on average more than 1 s longer per simulation. There is also a difference between BFAST and BFAST Monitor, with BFAST being on average much slower. EWMACD and BFAST Monitor performed similarly in terms of mean time per simulation but BFAST Monitor was less variable and slightly faster.

Table 3.3: Mean time to run per simulation set in seconds $\pm 1SD$. Numbers are rounded to 3dp but calculations were carried out on raw values.

Simulation Type	**BFAST**	**BFAST Monitor**	**CCDC**	**CCDC (CV)**	**EWMACD**
No change	0.274	0.033	0.130	1.781	0.012
Trend only	0.281	0.040	0.131	1.893	0.097
Break/trend	1.010	0.058	0.108	1.353	0.042
Amplitude change	0.753	0.045	0.127	1.563	0.050
LOS change	0.725	0.041	0.136	1.523	0.035
NOS change	0.149	0.043	0.140	1.309	0.054
Overall mean	0.534 ± 0.314	0.043 ± 0.007	0.129 ± 0.010	1.570 ± 0.211	0.048 ± 0.026

3.2.2 Overall Summary

3.2.2.1 Definition of Correct/False Trend Results for EWMACD

Since EWMACD does not include a trend term, it flags condition (trend) changes as breaks. Therefore, a correct result for EWMACD for the trend only set is defined as a result where EWMACD detected at least one break. A specific number of breaks was not used because how often EMWACD flags a trend as a change depends heavily on the parameters used and the steepness of the trend. For example, Figure 3.5A shows the results of an initial run of EWMACD on a time series with a trend of 0.001. Since the trend in the data is not accounted for, the fitted model deviates fairly quickly from the real data and a change is quickly flagged in June 2008. After re-initialising, EWMACD flagged another change in February 2013. Figure 3.5B shows the result for a time series with a trend of 0.002. Due to the steeper slope, EWMACD detected five breaks in this time series overall, the maximum possible given the two-year training period. For these reasons the results for the trend only set for EWMACD in Table 3.4 for false breaks were excluded since there is no definition of a false break for EWMACD for that set.

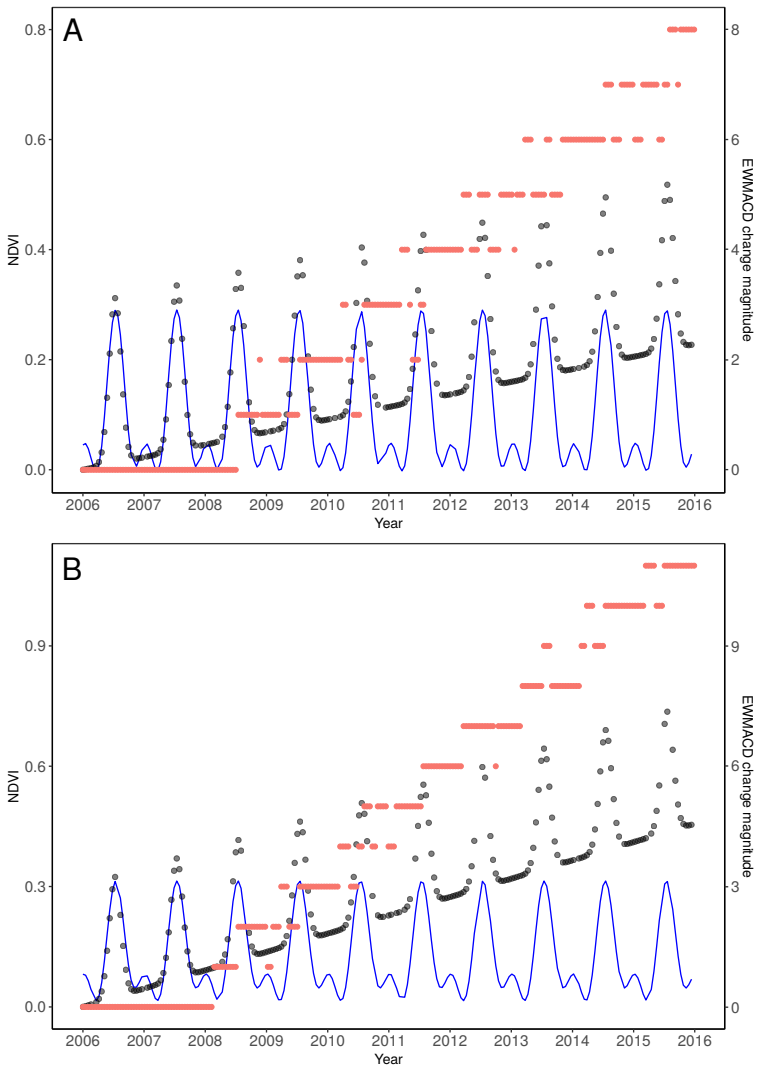


Figure 3.5: (**A**) Plot showing the results of an initial run of EWMACD on a time series with no noise, 10% missing data and a trend of 0.001. (**B**) Plot showing the results of an initial run of EWMACD on a time series with no noise, 10% missing data and a trend of 0.002. The magnitude of change as recorded in number of control limits is shown in red, where greater deviation from 0 indicates more deviation from the training period. The original time series values are shown in grey and the fitted seasonal model is shown in blue.

Additionally, for the break/trend set, breaks detected after the specified temporal window for correct break detection (96 days) are not counted as false breaks for EWMACD in cases where a trend greater than zero follows the break. While these constraints may produce a positive bias for EWMACD in terms of false break detection, this was considered to be the fairest way to maintain comparability between

EWMACD and the other methods as it does not require EWMACD to be parameterised differently for different simulation types.

Table 3.4: The percentage of correct and false results across all simulations. Correct breaks are those where either a break was detected within the specified temporal window (96 days for the break/trend set, one year for the seasonal change sets) or no break was detected where none existed. A correct result for EWMACD in the trend change set was defined as at least one break in trend being detected. False breaks are the percentage of results where either a break was detected where none existed or at least one break was detected outside of the specified temporal window. For EWMACD, if a trend was present in the data after the break then only changes detected before the true date of change were counted as false.

		Simulation set					
	Method	**None**	**Trend**	**Amplitude**	**LOS**	**NOS**	**Break/Trend**
Correct breaks (%)	BFAST	81.8	81.2	04.9	02.9	01.5	81.2
	BFASTM	40.4	40.4	53.7	28.8	71.3	57.8
	CCDC	97.5	97.6	14.6	09.3	32.9	64.1
	CCDC (CV)	87.3	96.3	20.3	13.5	51.5	65.8
	EWMACD	80.0	99.7	66.8	37.2	64.1	84.3
False breaks (%)	BFAST	18.3	18.8	53.3	39.4	05.5	29.2
	BFASTM	59.6	59.6	49.7	46.1	43.3	50.7
	CCDC	02.5	02.4	10.2	12.7	37.6	25.0
	CCDC (CV)	12.8	03.7	14.3	16.1	31.4	32.5
	EWMACD	20.0	00.3	39.9	39.5	50.2	17.6

3.2.2.2 True vs. False Changes

Table 3.4 provides an overall view of how each method performed on the different simulation sets. Results are presented as the percentage of simulations for that set for which the method either correctly identified there was no break (for the no change and trend only sets), or correctly identified a break within the specified temporal window (for the seasonal and break/trend sets). When considering Table 3.4, it is worth noting that, because all results are given as a percentage of the number of simulations in that set, the break/trend set contains 66% of all simulations and therefore performance in this set provides the best indication of overall method performance.

EWMACD gave the best performance in terms of overall effectiveness at break detection, correctly identifying a break where one existed in 76.6% of cases (i.e., excluding the no change and trend only sets). BFAST gave the second best performance (61.8%), followed by CCDC with CV (54.8%), BFAST Monitor (54.7%), and finally CCDC (51.8%). For false break detection (across all simulation sets), CCDC gave the best performance, detecting at least one false break in only 20.3% of simulations. For both the no change and trend only sets, CCDC correctly identified no break in nearly 100% of cases (Table 3.4). EWMACD gave the second best performance for false break detection overall (23.4%) and outperformed all other methods for the break/trend set (Table 3.4). However, performance on the three seasonal change sets was substantially worse than the other sets (Table 3.4). The distribution of false breaks for these sets was studied, where changes detected before the true date of change were counted as premature changes and changes detected more than one year after the true date of change were counted as late changes. For the amplitude change set, EWMACD detected at least one premature change in 8.1% of simulations and at least one late change in 33.8% of simulations. For change in LOS, the figures were 8.3% and 33.3%, respectively. For change in NOS, they were 17.9% and 37.1%, respectively.

CCDC with CV gave the third best performance (26.1%) and was slightly more likely to detect breaks in the no change and trend only sets than CCDC (Table 3.4). This was followed by BFAST which detected at least one false break in 30.5% of all simulations. However, while BFAST performed less well than CCDC and CCDC with CV in the no change and trend only sets, it still only detected breaks in those sets around 20% of the time.

BFAST Monitor gave the worst performance in terms of false breaks, detecting at least one false break in 50.9% of simulations. Table 3.4 shows that BFAST Monitor was more likely to detect a false break than any other method in four out of the six simulation sets and performed substantially worse than any other method on the no change and trend only sets. Given that the performance of BFAST Monitor on the no change and trend only sets was so poor, it was re-run on those sets using one harmonic term instead of two. Reducing the number of harmonics reduces the complexity of the fit, potentially leading to fewer false breaks. Using one harmonic did decrease the number of instances where at least one false break was detected to 42.7% for the no change set and 44.4% for the trend only set. However, using a single harmonic also decreased the percentage of correct breaks detected in the break/trend set from 57.8% to 51.1%, and increased the number of time series where at least one false break was detected from 50.7% to 53.9%.

BFAST Monitor outperformed all methods at identifying changes in NOS, and outperformed all methods except EWMACD at detecting changes in amplitude and LOS (Table 3.4). It found more correct breaks for the amplitude and change in LOS sets when the magnitude of the change was greater. For example, it detected 68.9% and 70.2% of breaks correctly for the 0.3 and -0.3 amplitude change values, respectively, but only 36.7% and 32.9% for the 0.1 and -0.1 change values. EWMACD showed a similar pattern, correctly identifying more than 70% of changes in amplitude for the 0.3 and -0.3 levels but less than 40% of changes for the 0.1 and -0.1 levels.

For change in LOS, EWMACD detected only 5.0% of breaks correctly for $\Delta c_1 = 5$ while BFAST Monitor performed slightly better at 9.5%. Performance for $\Delta c_1 = 30$ was much better for both EWMACD (65.4%) and BFAST Monitor (51.2%).

In contrast to EWMACD and BFAST Monitor, the other three methods performed poorly on the seasonal change sets. BFAST consistently failed to detect seasonal breaks of any type. It was the least effective method for detecting the onset of changes in amplitude, LOS, or NOS, but did frequently report at least one false change in those simulation sets (Table 3.4). CCDC and CCDC with CV were both more likely to detect a correct change for the change in NOS set than for the other two sets, where they performed similarly to BFAST (Table 3.4). However, CCDC and CCDC with CV were more likely to detect at least one false break in the change in NOS set than in the change in amplitude or change in LOS sets, whereas for BFAST the opposite was true. In the case of the change in amplitude set, BFAST detected at least one false break more than 50% of the time (Table 3.4), higher than any other method.

3.2.3 By Noise Level

Figure 3.6 shows a breakdown of the results from Table 3.4 by noise level. RMSE number of breaks is also included here because it provides an idea of whether methods tend to detect more or less breaks overall given increasing levels of noise, regardless of whether those breaks are correct or false. A method which always detected one break would have an RMSE of zero; however, the method could still be poor at estimating the timing of the break.

Results for the trend only set for EWMACD are not included in the breakdown plots of noise/missing data. Instead, percentages reflect correct results/false breaks found within the remaining simulation sets. This is because there is no way to include the trend only set in the false break and RMSE number of breaks plots for EWMACD

because there is no way to determine error. Given that the definition of a correct break for the trend only set for EWMACD is more lenient than for other methods, and that EWMACD correctly identified a break nearly 100% of the time, the trend only set was also excluded from the correct breaks plots in order to create a fairer comparison.

All methods reported a significant negative correlation between percentage of correct results found (across all simulations) and noise level ($p < 0.01$, $\rho < -0.9$). The decrease in percentage of correct results found between the lowest and highest noise levels was approximately 20% less for BFAST than for any other method (Figure 3.6), suggesting more consistent performance across noise levels in this metric than the other methods.

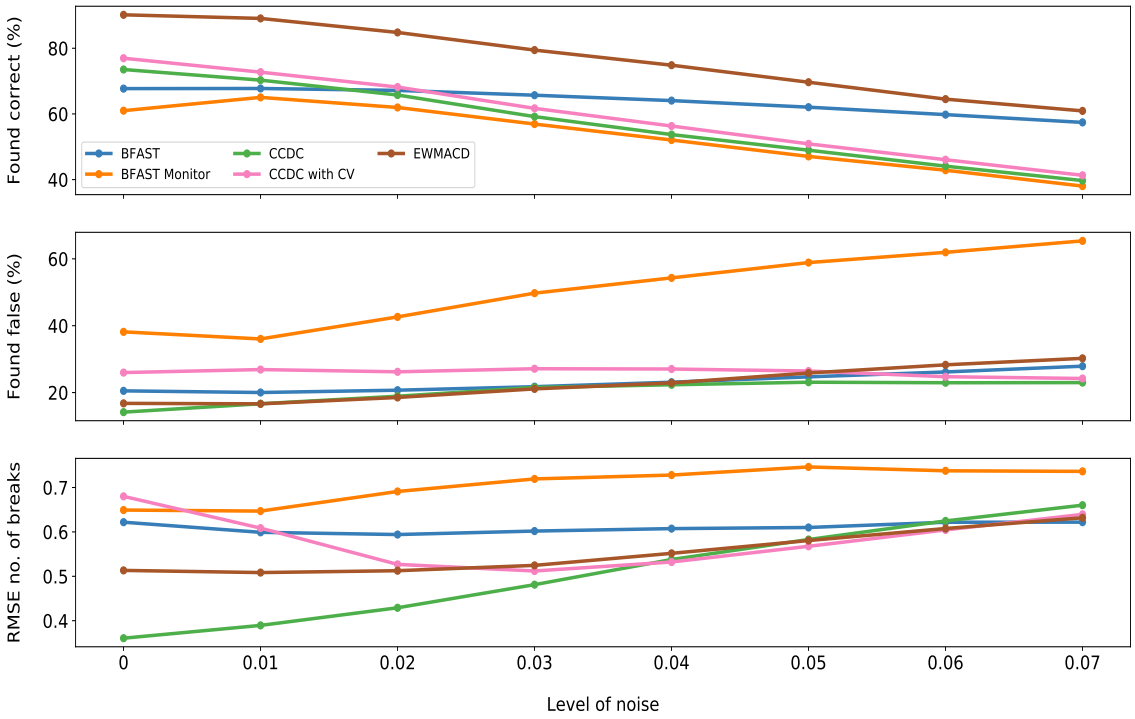


Figure 3.6: Plots showing the percentage of simulations with a correct result, the percentage of simulations where at least one false break was detected, and the RMSE number of breaks, per level of noise. Correct results include simulations where absence of change was correctly identified. Results for EWMACD for the trend only set are not included (with n adjusted accordingly) because EWMACD was not tuned to detect a specific number of breaks in that set and therefore number of false breaks and RMSE could not be calculated.

All methods apart from CCDC with CV also reported significant positive correlations between noise level and percentage of results where at least one false break was found ($p < 0.01$, $\rho > 0.9$). Generally, the results for this metric are very similar for CCDC, BFAST, and EWMACD. The results for BFAST Monitor follow a more extreme trend, increasing from 38.2% to 65.4%. At higher levels of noise BFAST Monitor was substantially more likely to detect at least one false break in a time series than any other method (Figure 3.6).

No significant correlation was found between noise level and RMSE number of breaks for BFAST or CCDC with CV. Figure 3.6 indicates that while BFAST did not

have the lowest RMSE values, it remained very consistent across noise levels. While a significant positive relationship was reported for EWMACD ($p < 0.01$, $\rho = 0.93$), it was also relatively consistent across noise levels for RMSE.

Above a noise level of 0.03, RMSE number of breaks for EWMACD, CCDC, and CCDC with CV is very similar. However, CCDC with CV shows a complex relationship between RMSE number of breaks and noise whereby it is more likely to detect the correct number of breaks with either very low or very high noise levels (Figure 3.6). In contrast, RMSE for CCDC with a fixed λ increased substantially with noise, from 0.36 to 0.66; the largest increase of any method. This relationship was significant ($p < 0.01$, $\rho = 1.00$). Below a noise level of 0.03, CCDC reported the lowest RMSE for number of breaks of any method, indicating that at low noise levels it is the most likely to correctly estimate the number of breaks in a time series. There was therefore a large difference in RMSE number of breaks between CCDC and CCDC with CV at the lowest noise levels. A significant positive relationship was also found between RMSE number of breaks and noise level for BFAST Monitor ($p < 0.01$, $\rho = 0.88$), which performed less well than any other method except for at the lowest noise level where CCDC with CV was worse.

3.2.4 By Missing Data Level

Figure 3.7 shows a breakdown of the results in Table 3.4 by percentage of data missing. As with Figure 3.6, RMSE number of breaks is included here because it provides an idea of whether methods tend to detect more or less breaks overall given increasing levels of missing data, regardless of whether those breaks are correct or false.

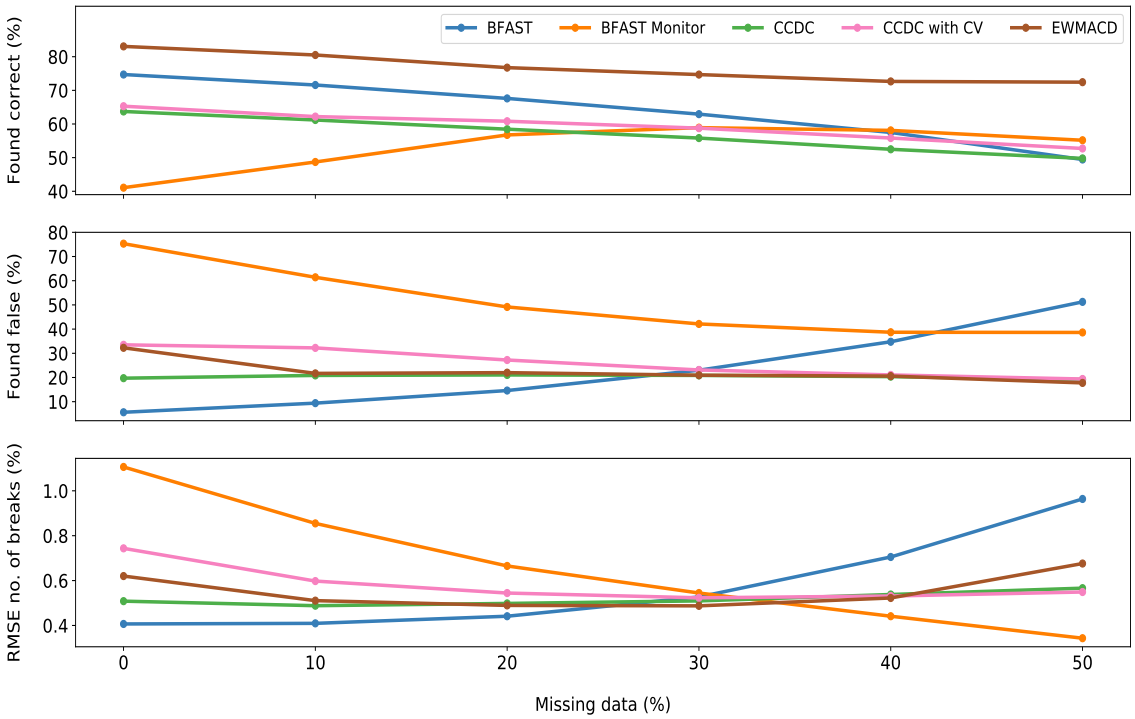


Figure 3.7: Plots showing the percentage of simulations with a correct result, the percentage of simulations where at least one false break was detected, and the RMSE number of breaks, per level of missing data. Correct results include simulations where absence of change was correctly identified. Results for EWMACD for the trend only set are not included (with n adjusted accordingly) because EWMACD was not tuned to detect a specific number of breaks in that set and therefore the number of false breaks and RMSE could not be calculated.

Breaking down the results by percentage of missing data revealed contrasting trends for BFAST and BFAST Monitor. For BFAST, significant ($p < 0.01$) positive correlations ($\rho = 1.00$) were found between level of missing data and percentage of simulations where at least one false break was found and between missing data level and RMSE number of breaks. However, this was reversed for BFAST Monitor, where significant negative trends ($p < 0.01$, $\rho = -1.00$) were found for both metrics. A significant negative correlation ($p < 0.01$, $\rho = -1.00$) was also found for BFAST between missing data level and percentage of simulations where the correct break was found, whereas no significant trend was found for BFAST Monitor. Figure

3.7 shows an increasing trend for BFAST Monitor in terms of correct breaks found up to the 30% level, and then a slight decreasing trend. Overall, the results show that BFAST becomes less effective at break detection with more missing data, while BFAST Monitor becomes more effective.

CCDC with a fixed λ and CCDC with CV performed very similarly overall. Along with EWMACD, performance for these methods was more consistent than for BFAST and BFAST Monitor across missing data levels. Figure 3.7 indicates that, while CCDC was generally less likely to identify the correct break with increasing missing data, there was little effect of missing data level on the percentage of simulations where at least one false break was found or on RMSE number of breaks. This is confirmed by Spearman's rank tests, which indicate a negative correlation of missing data level with percentage of correct breaks found ($p < 0.01$, $\rho = -1.00$) but no significant correlation of missing data level with the other two metrics. CCDC with CV also showed no significant correlation of missing data level with RMSE number of breaks and a significant negative correlation between missing data level and percentage of correct breaks found ($p < 0.01$, $\rho = -1.00$). However, Figure 3.7 indicates that CCDC with CV was more likely than CCDC to overestimate number of breaks for levels below 30%. Unlike CCDC, CCDC with CV showed a significant negative correlation between missing data level and percentage of results with at least one false break ($p < 0.01$, $\rho = -1.00$). Figure 3.7 indicates that this is due to CCDC with CV being more likely than CCDC to detect at least one false break for missing data levels below 40%.

As with CCDC and CCDC with CV, a Spearman's test showed no significant correlation for EWMACD between missing data level RMSE number of breaks. Figure 3.7 shows that this is because EWMACD was better at estimating the number of breaks at very high and very low levels of missing data. There were significant negative correlations with percentage of true breaks detected and percentage of results where

at least one false break was found ($p < 0.01$, $\rho < -0.9$). While both EWMACD and BFAST Monitor showed significant negative correlations between missing data level and percentage of simulations where at least one false break was found, the trend was much less pronounced for EWMACD (Figure 3.7). While a significant trend was found for percentage of correct breaks detected, EWMACD appears generally less affected by missing data for this metric than any other method (Figure 3.7).

3.2.5 Break Magnitude by Noise and Missing Data

Figure 3.8 shows RMSE break magnitude for each method by level of noise and by level of missing data, for all correctly identified breaks in the break/trend set. RMSE break magnitude was not investigated for the seasonal change sets because no method except BFAST was designed to estimate the magnitude of seasonal changes. Forty-three data points were removed from this dataset for BFAST Monitor because the estimated break magnitude for those breaks was extremely unrealistic and outside the possible range for a change in NDVI (i.e., a maximum change magnitude of ± 2). Given that the size of the dataset is 100,800 simulations, this represents a very small proportion of the data. Without these outliers, the RMSE break magnitude results are likely to be much closer to typical performance for BFAST Monitor.

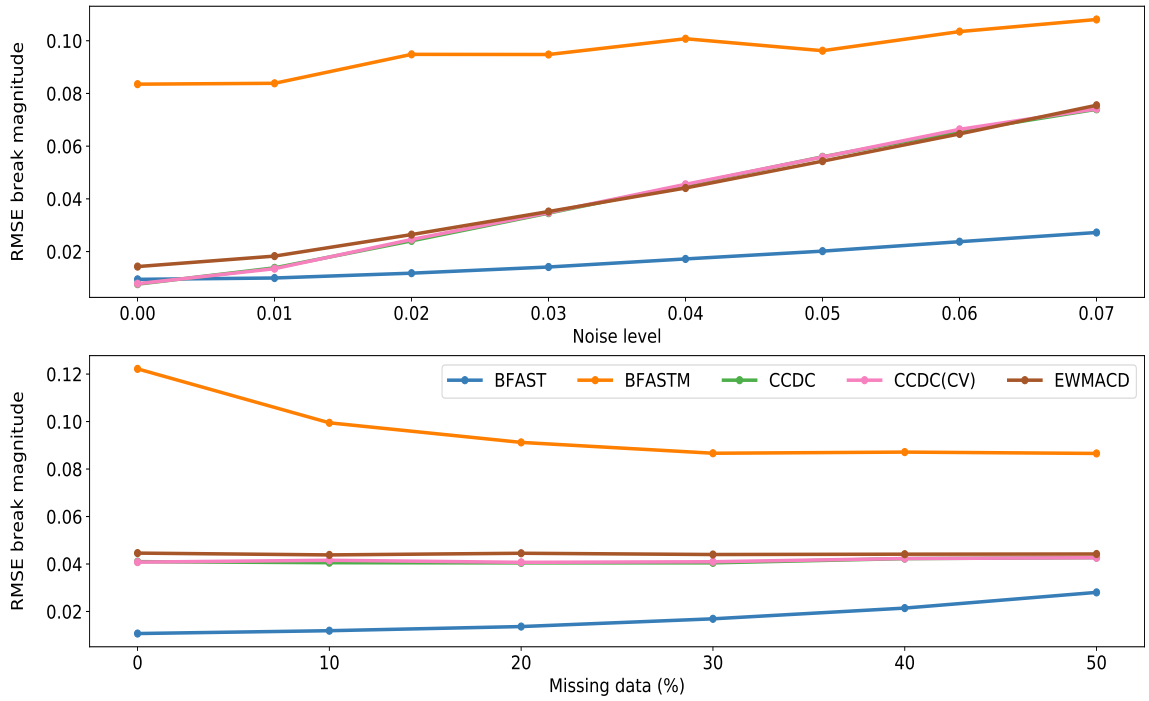


Figure 3.8: Plots showing RMSE break size vs. percentage of missing data and RMSE break size vs. noise level for all correctly detected changes in the break/trend set. A correct change is defined as a change found no more than 96 days after the true date of change.

All methods showed a significant positive correlation between noise level and RMSE break magnitude ($p < 0.01$, $\rho > 0.95$) using a Spearman's rank correlation test. CCDC, CCDC with CV, and EWMACD showed no significant correlation between missing data level and RMSE break magnitude. However, BFAST and BFAST Monitor reported significant ($p < 0.01$) positive ($\rho = 1.00$) and negative ($\rho = -0.94$) trends, respectively. It is clear in Figure 3.8 that CCDC and CCDC with CV performed almost identically at estimating break magnitude across all noise and missing data levels. EWMACD also produced a very similar result to the CCDC methods. BFAST consistently performed better than any other method, and BFAST Monitor consistently performed worse.

3.2.6 By Break Severity

To more closely investigate the ability of the different methods to detect different types of break, the break/trend set was broken down into three categories of break: Extreme, Moderate, and Subtle. This was not done for the seasonal change sets (change in amplitude, change in LOS, and change in NOS) because, as Table 3.4 indicates, results for those sets were generally poor. Simulations were categorised based on level of break and level of trend following the break, with the idea being that larger breaks followed by strong positive or negative trends are easier to detect than smaller magnitude breaks followed by weak trends or no trend at all. Break severity was therefore categorised as follows:

- **Extreme breaks** have a large or medium magnitude break (break > 0.1 or break < -0.1) followed by a strong or medium trend (trend > 0.001 or trend < -0.001). $n = 38,400$.

- **Moderate breaks** have a large break (break $= 0.3$ or break $= -0.3$) with a weak trend (trend $= 0.001$ or trend $= -0.001$), a large break with no trend, a small break (break $= 0.1$ or break $= -0.1$) with a strong trend (trend $= 0.002$ or trend $= -0.002$), or a medium break (break $= 0.2$ or break $= -0.2$) with a weak trend (trend $= 0.001$ or trend $= -0.001$). $n = 33,600$.

- **Subtle breaks** have a small break (break $= 0.1$ or break $= -0.1$) with a weak trend (trend $= 0.001$ or trend $= -0.001$), a small break with no trend, a small break with a medium trend (trend $= 0.0015$ or trend $= -0.0015$), or a medium break (break $= 0.2$ or break $= -0.2$) with no trend. $n = 28,800$.

As described in Section 3.2.5, when calculating RMSE break magnitude, 43 data points were removed from the dataset for BFAST Monitor. Table 3.5 shows all methods were less likely to detect the correct break in a time series as break severity

decreased. The largest decreases were for CCDC and CCDC with CV, with differences of 44.4% and 42.2%, respectively, between the Extreme and Subtle simulation sets. In contrast, the reduction between these sets for EWMACD was around four times less at 10.5%.

Table 3.5: Table showing the percentage of results where the correct break was identified, the percentage of results where at least one false break was found, and the RMSE estimated break magnitude for correctly detected breaks, for the break/trend set. Correctly identified changes are those detected no more than 96 days after the true date of change. Extreme changes are those with a large or medium magnitude break followed by a strong or medium trend. Moderate changes have a large break followed by a weak trend or no trend, a small break followed by a strong trend, or a medium break followed by a weak trend. Subtle changes have a small break followed by no trend, a weak trend, or a medium trend, or a medium break followed by no trend.

	Method	**Extreme**	**Moderate**	**Subtle**
Correct breaks (%)	BFAST	85.8	82.0	74.2
	BFAST Monitor	67.1	59.6	43.2
	CCDC	81.3	67.8	36.9
	CCDC with CV	78.8	68.2	45.8
	EWMACD	88.4	84.9	77.9
False breaks (%)	BFAST	25.7	28.5	34.5
	BFAST Monitor	43.3	49.4	61.9
	CCDC	18.1	26.3	32.8
	CCDC with CV	27.9	34.1	36.6
	EWMACD	13.5	15.1	20.4
RMSE magnitude	BFAST	0.02	0.02	0.02
	BFAST Monitor	0.12	0.08	0.06
	CCDC	0.04	0.04	0.03
	CCDC with CV	0.04	0.04	0.03
	EWMACD	0.04	0.04	0.04

All methods were more likely to detect at least one false break in a time series as break severity decreased (Table 3.5). The method with the largest increase between the Extreme and Subtle simulation sets was BFAST Monitor (18.6%). The method with the smallest increase was CCDC with CV (8.7%).

Table 3.5 also shows that, when considering breaks which were correctly identified, BFAST Monitor was 50% better at estimating the magnitude of Subtle breaks than

Extreme breaks. For all other methods, the change in RMSE magnitude between the Extreme and Subtle change sets was 0.01 or less.

3.3 Discussion

3.3.1 BFAST

BFAST is a widely used method and a recent study by Saxena et al. [263] found that it rarely failed to detect breaks in time series. It was found found that, for simulations with a change, BFAST correctly identified more changes than any method other than EWMACD. It was also the most accurate method for estimating the magnitude of breaks and performed fairly consistently across noise levels. BFAST estimates break magnitude using the models fitted both before and after the break, which is more difficult for live monitoring methods which cannot fit a new stable model until multiple new observations are available. Given that BFAST receives the whole time series at once, it is also not unexpected that it was found to be more robust to noise than the live monitoring methods, which are more likely to be influenced by a single noisy data point. BFAST also performed relatively consistently across a range of different change severity levels for the break/trend set.

Given that BFAST uses an iterative process to find breaks, it is not unexpected that it was found to be slower than other methods at processing time series. BFAST was faster when applied to the simulations with no real changes. This was probably because BFAST first evaluates the possibility of any change being present using the OLS-MOSUM test. Optimisation of breakpoints is only carried out if the OLS-MOSUM test indicates a structural change within the time series.

BFAST appeared to be more affected by missing data than other methods. Unlike all other methods, BFAST was more likely to detect at least one false break the more data were missing and more likely to incorrectly estimate the number of breaks overall. A possible reason for this is the linear interpolation used to create a daily

time series for processing with BFAST. The more data are removed, the less well this interpolation will represent the true temporal trajectory of the data. This is one possible explanation for BFAST's high likelihood of detecting at least one false break in this study.

BFAST is the only method used in this study which explicitly aims to detect seasonal changes separately from trend changes [296]. However, BFAST performed very poorly in this study at detecting seasonal changes such as a change in amplitude, change in LOS, or change in the number of seasons. This poor performance existed across all magnitudes of change for the change in amplitude and change in LOS sets. A possible explanation is that these changes are too easily accounted for by the trend component. Given the whole time series at once, BFAST attempts to fit the optimal number of breakpoints to both the trend and seasonal component. However, as seen in Figure 3.9A, a decrease in signal amplitude results in an overall decrease in NDVI and can be interpreted as a break in trend. In Figure 3.9B, the change in LOS has been interpreted as a trend across the time series and no break is detected. The tendency of BFAST to account for amplitude and LOS changes by assuming a steeper trend is probably why BFAST was more likely to report at least one false break in this simulation set (Table 3.4), since trend breaks were not counted as correct for seasonal breaks even if they were temporally correct.

BFAST detected very few changes in NOS correctly. Figure 3.9C shows that BFAST could, on some occasions, correctly detect this type of change. However, in some cases, BFAST simply fitted a more complex seasonal model to the entire time series. The effect of this can be seen in Figure 3.9D.

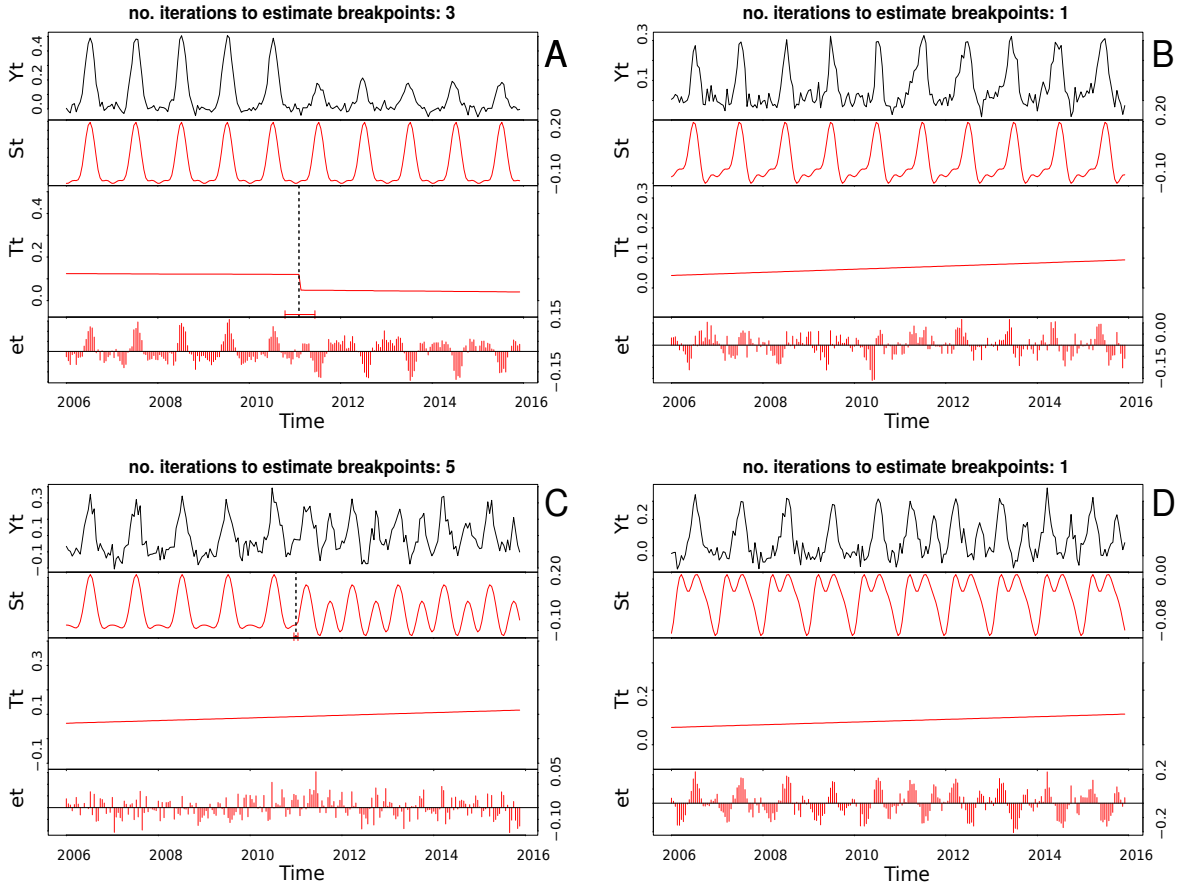


Figure 3.9: Output from the BFAST R package for: (**A**) time series with a change in amplitude of 0.3, a noise level of 0.02, and 10% missing data; (**B**) time series with a change in SOS of −37 days, a noise level of 0.03, and no missing data; (**C**) time series with a change from one seasons to two, a noise level of 0.04, and 10% missing data; and (**D**) Time series with a change from one seasons to two, a noise level of 0.03, and no missing data. *Yt*, original signal; *St*, decomposed seasonal component; *Tt*, decomposed trend component; *et*, error. The plots presented are direct outputs of the R BFAST package.

3.3.2 BFAST Monitor

BFAST Monitor was the fastest method and the most consistent in average runtime across simulation types. Overall, it came second to last at correctly identifying breaks in time series where they existed; only CCDC performed worse. BFAST Monitor also consistently detected more breaks than were present in the time series. This may explain to some extent why BFAST Monitor was so poor at finding true breaks; if a false break is detected in the two years before the true break, the true break is likely

to be missed when re-initialising with a new stable history period.

It was considered that BFAST Monitor's high false break detection rate might be a result of the second-order harmonic model overfitting the data. Previous studies have used a single harmonic model with BFAST Monitor in areas with low observation frequency where the underlying data were known to follow a simple seasonal curve [64, 65]. However, while using a simpler model did reduce the number of time series where at least one false break was detected for the no change and trend only sets, it increased it for the break/trend set. This indicates that a single-order harmonic model was too simple for the underlying data. The second-order harmonic used for BFAST Monitor was also the same order as that used for EWMACD, which was far less likely to detect at least one false break for the no change set. Given that EWMACD does not incorporate a trend term, it is possible that BFAST Monitor simply has more dimensions in which to overfit.

BFAST Monitor performed similarly to EWMACD on the seasonal change sets and better than BFAST, CCDC, or CCDC with CV. Given that BFAST Monitor was better at detecting larger magnitude changes for the amplitude and change in LOS sets, there is evidence that it can correctly identify seasonal changes, especially if they are large. However, BFAST Monitor performed worst overall at estimating the number of breaks in a time series. Alongside the high false break detection rates, this suggests that often even when BFAST Monitor does detect a break correctly, it will detect other non-existent breaks in the same time series.

Interestingly, BFAST Monitor was the only method which improved substantially the more data were missing. While this trend is less pronounced for percentage of correct breaks detected, it is clear that BFAST Monitor becomes both less likely to detect at least one false break and more likely to correctly estimate the number of breaks present with increased missing data (Figure 3.7). This is contrary to what would generally be believed, i.e. that more data equal better break estimation.

Given this result, it could be concluded that BFAST Monitor is preferable to the other methods in regions with high quantities of missing data, e.g. regions with high cloud cover. However, given its high rate of false break detection, removing data may simply remove opportunities for breakpoints since BFAST Monitor operates on an observation-by-observation basis. Unlike the other live monitoring methods, the MOSUM method used by BFAST Monitor only requires a single observation to exceed a boundary for a change to be flagged [297, 328]. While this allows for faster detection of breaks, it can lead to far more observations being flagged as changes.

In terms of estimating break magnitude, BFAST Monitor performed poorly, being the least accurate of all the methods (Figure 3.8). However, there was around a 50% improvement in RMSE break magnitude between the Extreme and Subtle change sets for the break/trend simulation set, suggesting that BFAST Monitor struggles to accurately estimate larger breaks. This is possibly because the method used by BFAST Monitor to estimate break magnitude is based on median values. For breaks followed by strong trends, this could result in break size being underestimated as in Figure 3.10A, or overestimated as in Figure 3.10B, because the trend component causes the median of the trend segment to move closer to or further away from the values of the first segment. It could be argued that BFAST Monitor is actually providing more information about the change here because break magnitude is influenced by the direction of recovery. The usefulness of that additional information will depend on the intention of the study being undertaken.

3.3.3 CCDC

CCDC ranked in the middle for runtime, being much faster than CCDC with CV and BFAST but slower than BFAST Monitor and EWMACD. Overall, CCDC detected at least one false break in the fewest number of time series, but was worse than any other method at identifying changes where they existed. Based on this study,

the main strength of CCDC is that it is unlikely to overestimate the number of breaks. However, this comes at the cost of also being more likely to miss breaks where they exist. It must be borne in mind that the purpose of CCDC is to detect complete changes in land cover type. While the classification element of CCDC was not discussed here, its lack of sensitivity to smaller changes might be a positive in this regard. The emphasis on changes in class is also probably why CCDC performed so poorly at detecting seasonal changes. A change in land cover class is likely to result in more complex changes to the shape of the seasonal curve than a straightforward change in amplitude or LOS.

There is evidence for this in the observation that CCDC did detect more breaks, both correct and false, in the NOS change set than in the change in amplitude or change in LOS sets. Since CCDC uses the RMSE of models to find changes, if a model underfits the seasonal curve, then many seasonal changes may not exceed the six times RMSE level required for CCDC to confidently flag a change. Figures 3.11A and 3.12A show that CCDC sometimes failed to properly capture the amplitude and shape of time series. A change in the number of seasons introduces changes at the start, end, and middle of the season; multiple points at which the previous model can fail to fit.

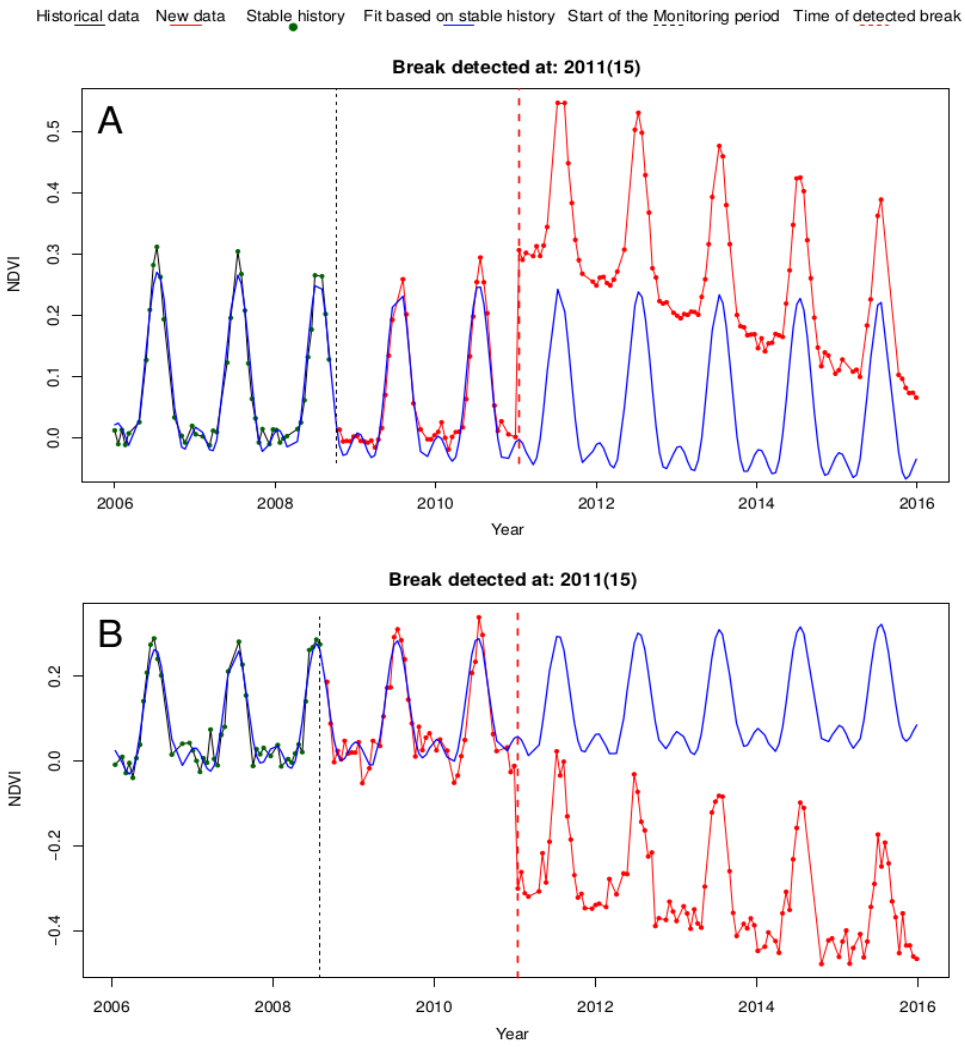


Figure 3.10: (**A**) Result for BFAST Monitor for a time series with an abrupt change of 0.3 followed by a trend of −0.002, a noise level of 0.01, and 20% missing data. Break magnitude estimated by BFAST Monitor was 0.17. (**B**) Result for BFAST Monitor for a time series with an abrupt change of −0.3 followed by a trend of −0.0015, a noise level of 0.03, and 20% missing data. Break magnitude estimated by BFAST Monitor was −0.38. The plots presented are direct outputs of the R BFAST Monitor package.

It was considered that the tendency of CCDC to underfit was due to the selected value of $\lambda = 0.01$. However, CCDC with CV did not perform substantially differently to CCDC, leading us to believe that the Lasso fitting method is generally more likely to underfit the seasonal curve than OLS fitting. This makes it a suitable choice if the aim is to detect only the more substantial changes in a time series. CCDC also detected around 50% fewer breaks in the Subtle change category compared to

the Extreme change category for the break/trend set, suggesting that CCDC is also unlikely to detect more minor breaks or trend changes which again could be associated to within-class rather than between-class changes. However, CCDC with fixed λ may provide more opportunity to control the degree of over- or underfitting. Given the increase in speed over using cross validation, using a fixed value for λ is therefore preferable if the value is chosen carefully.

CCDC became substantially worse at estimating the number of breaks and detecting correct breaks (or correct absence of breaks) with increasing noise level, although percentage of simulations where at least one false break was detected did not increase (Figure 3.6). This suggests that, as noise increases, CCDC is more likely to miss breaks altogether rather than attributing an incorrect date of change. Both percentage of results where at least one false break was found and RMSE number of breaks for CCDC were stable across missing data levels. Therefore, the evidence is that noisy data are more likely to affect the efficacy of CCDC in correctly identifying breaks than missing data. This supports the conclusion that CCDC is more suited to situations requiring robust identification of complete changes in land cover, as it is unlikely to flag smaller changes in noisy time series.

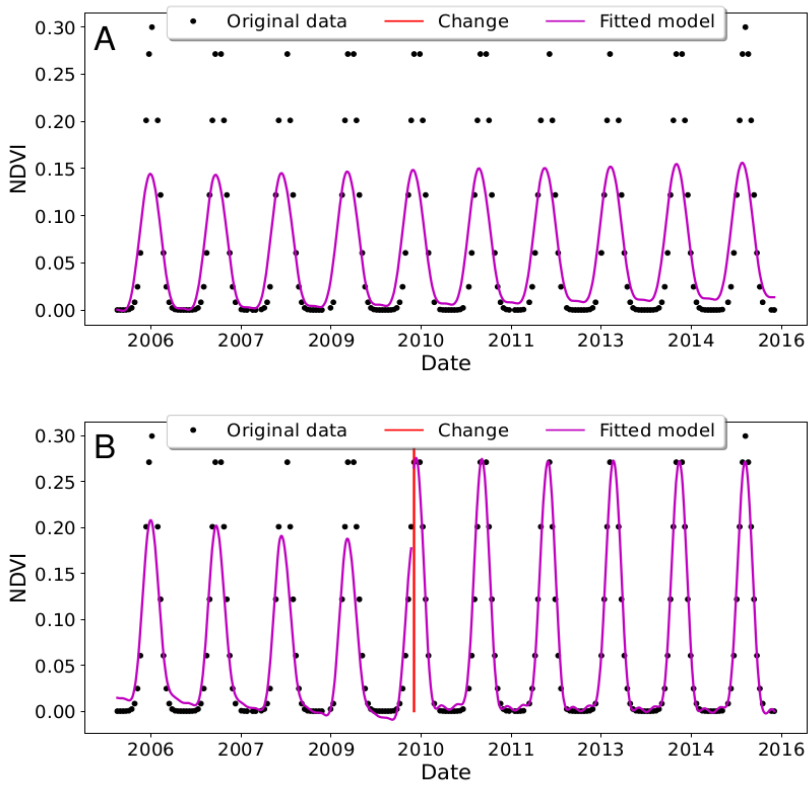


Figure 3.11: Plots showing output from: (**A**) CCDC with a fixed λ; and (**B**) CCDC with CV for a time series with no changes, no noise, and 10% missing data.

CCDC, CCDC with CV, and EWMACD all performed very similarly at estimating break magnitude; this was expected given that the same method was used for all three. In general, this method of break estimation appears to be robust to missing data but does get less effective with increased noise. The effect of noise is not surprising given that this method relies on residual values; the more noisy the data are, the less likely that value is to reflect the true break size. The method used by BFAST had much lower RMSE and was more robust against noise.

3.3.4 CCDC with CV

The purpose of using a cross-validated approach was to investigate whether allowing λ to vary would produce substantial improvement over a fixed λ. Not surprisingly given that cross-validation requires fitting large numbers of models, CCDC with CV

took much longer to run than the other methods.

Overall, it was found that using CV made CCDC more likely to detect true breaks, but also more likely to detect at least one false break in a time series. This suggests that using a cross-validated approach did lead to more closely fitting models than using $\lambda = 0.01$ and to overfitting in some cases. Figure 3.12 shows an example where CCDC with CV detects two additional breaks where CCDC estimates the number of breaks correctly.

One observation made was that, unlike CCDC, CCDC with CV did not have a straightforward relationship between RMSE number of breaks and noise. CCDC with CV was found to be less accurate at detecting the number of breaks at the lowest and highest noise levels than at the intermediate levels. With increased noise, the method was less likely to detect correct results and the likelihood of detecting at least one false break remained constant. However, the change in RMSE tells us that the actual number of false breaks found is likely to be higher at extremes of noise. At high levels of noise, models are more likely to be influenced by noisy data points and may be fitting to noise. This means that more false breaks get detected, but fewer true breaks. This is probably why most of the methods performed less well in terms of RMSE number of breaks at high noise levels. The unique pattern shown by CCDC with CV suggests that it must also be detecting more breaks if there is very little noise. With less noise, CCDC with CV may be fitting the data too closely, leading to more false breaks per simulation.

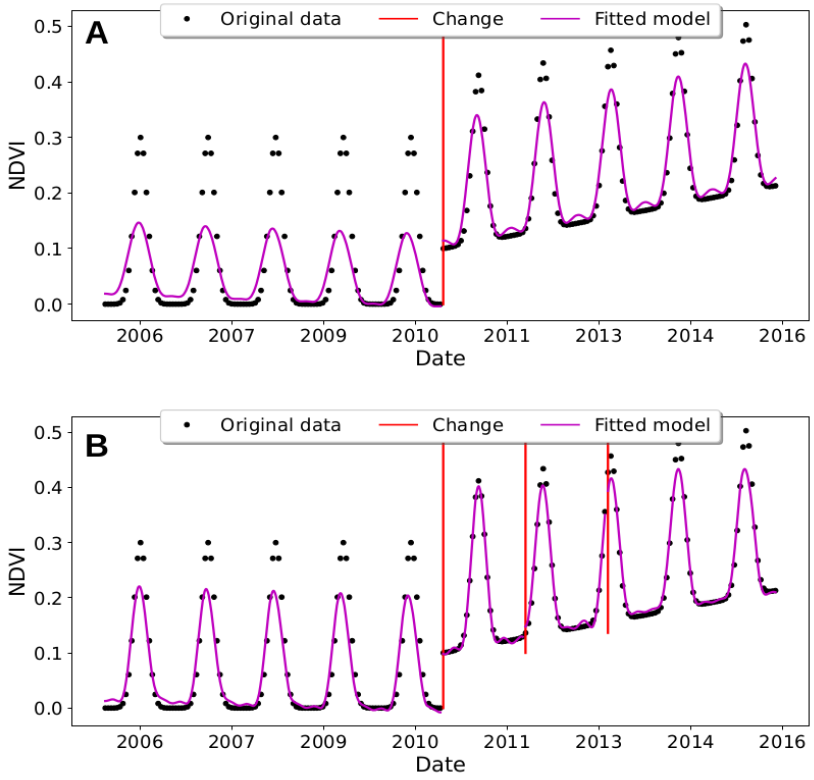


Figure 3.12: Plots showing output from: (**A**) CCDC with a fixed λ; and (**B**) CCDC with CV for a time series with an abrupt change of 0.1 followed by a trend of 0.001, with no noise and no missing data.

CCDC with CV was slightly more likely to overestimate the number of breaks and more likely to detect at least one false break in a time series than CCDC at missing data levels less than 40%. These trends are much less pronounced than for BFAST Monitor, and since CCDC requires six observations to confidently flag a change, the cause is likely to be different. In the case of CCDC with CV, this effect is probably again due to a tendency to overfit; with fewer data, CCDC with CV has fewer points to fit to. Figure 3.11B shows the output from CCDC with CV for a time series with no breaks and very few missing data, where the algorithm detects a non-existent break.

CCDC with CV performed similarly to CCDC in the breakdown by change severity. One notable difference is that CCDC with CV was more likely to correctly

identify breaks in the Subtle change category, and was more likely overall to detect at least one false break in a time series. This reinforces the previous point that while in general CCDC is not designed to detect lower magnitude changes, some control over sensitivity can be gained by setting the value of λ appropriately. Lasso fitting attempts to regularise the model coefficients, in some cases reducing them to zero, resulting in a form of feature selection. Larger values of λ will increase the degree of regularisation and therefore increase the likelihood that some seasonal coefficients will be reduced to zero. CCDC was never intended to be able to detect seasonal breaks [340, 342, 344] and therefore it is not unexpected that both CCDC variants performed poorly on these simulation sets.

3.3.5 EWMACD

EWMACD is designed to detect subtler changes, such as partial disturbance of forest pixels [35]. This claim is supported by the results of the simulation testing. Across all simulations with a change, EWMACD was by far the most effective at correctly identifying the date of change. The high overall rate of correct change estimation is due to EWMACD's good performance across both the break/trend and seasonal change sets. While EWMACD was not able to detect all seasonal changes, overall it outperformed any other method. For higher magnitudes of seasonal change, EWMACD performed very well at correctly identifying changes. Figure 3.13A shows how a change in amplitude results in deviation from the history period, causing deviation in the control chart around the peaks of the seasonal curves.

Across all simulations, only CCDC was less likely to detect at least one false break. However, EWMACD was much more likely to detect false changes in seasonal sets than in other sets. It must be remembered that, for the break/trend set, any changes after the 96-day window were not counted as false for EWMACD because it is designed to detect trend changes. The lower likelihood of detecting at least one

break in a time series for the break/trend set is probably because of this. However, EWMACD only detected a change in the no change set around 20% of the time. It was also found that EWMACD was around three times more likely to detect a false break after the date of true change as before. This suggests that the higher rate of false change detection in the seasonal change sets might partly be caused by changes being detected too late, as shown in Figure 3.13B. It is also possible that the vertex method used to find the next stable period after a change does not work as well for seasonal changes where the historic model still fits some parts of the seasonal curve.

EWMACD's response to noise was as expected in that, the noisier are the data, the less likely EWMACD was to correctly identify the number and location of breaks. At noise levels less than 0.03, only CCDC performed better in terms of RMSE number of breaks. At high noise levels, EWMACD was still more likely to correctly identify a break (or lack of one) than any other method. As with CCDC and CCDC with CV, break magnitude estimation got worse with noise, but not with missing data, as discussed in previous sections.

EWMACD did show an unusual result for response of RMSE number of breaks to missing data level, whereby RMSE was higher for the 0 and 50% levels than for the levels in between. EWMACD was also more likely to detect at least one false break with no missing data than at any other level, although this evened out over the remaining levels. Percentage of correct results had an overall downward trend. This suggests that, when there are no missing data, EWMACD is detecting too many changes, whereas, at the 50% missing data level, it is detecting too few. This is possibly because if more data are available, EWMACD is more likely to overfit the data than the CCDC methods, which are closest to EWMACD in trend. With 50% of the data missing, EWMACD may reach a tipping point where it lacks enough data to adequately fit the model, leading to a less well fitting model and making change detection more difficult.

EWMACD had the second fastest runtime behind BFAST Monitor, although runtime was more variable than for CCDC or for BFAST Monitor. The increased variability was partly because runtime on the trend only set was higher than the other sets due to EWMACD detecting more changes and therefore needing to output more results. Given that EWMACD performed far better at break detection than BFAST Monitor it can be concluded that for seasonal breaks and subtler abrupt changes it is the preferred method.

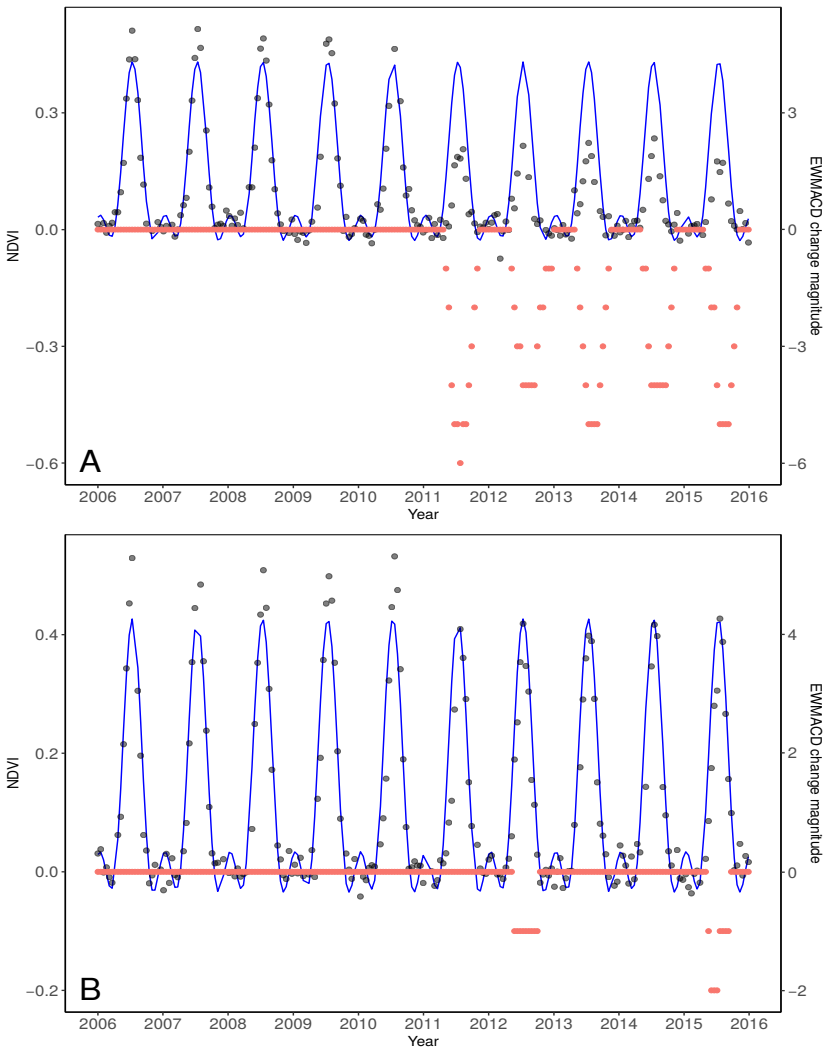


Figure 3.13: (**A**) Output from EWMACD for a time series with a change in amplitude of −0.3, a noise level of 0.02 and 10% missing data where the date of change was correctly identified. (**B**) Output from EWMACD for a time series with a change in amplitude of −0.1, a noise level of 0.02 and 10% missing data. EWMACD first detects a break on the 9 May 2012, after the correct date of change. The magnitude of change as recorded in number of control limits is shown in red, where greater deviation from 0 indicates more deviation from the training period. The original time series values are shown in grey and the fitted seasonal model is shown in blue.

3.3.6 Limitations

It is recognised that using simulated data in place of real-world observations has its limitations. The simulations used in this study are essentially idealised time series as it is very difficult to realistically simulate the levels of variation that exist in the real world. This will bias results towards being over-optimistic, and all of the

methods studied will likely perform less well on real-world data. However, very little optimisation was carried out in an attempt to keep the methods as comparable as possible. Some result are therefore likely to be negatively biased in comparison to more specific real-world applications. To facilitate comparability and examine off-the-shelf performance, this study used parameters which were not optimised for the the presented problems. In many cases, real data will also contain multiple breakpoints; however, assuming a sufficient stable period between changes, accuracy would likely be very similar across all breakpoints.

3.3.7 Future Work

This study presents evidence that simulated data enable a robust way of evaluating different change detection approaches under a range of scenarios, which would not be possible using real data. Using simulations can provide a benchmark against which to test new methods as well as an objective way to compare different methods and determine their strengths and weaknesses. Simulations could also be used when proposing new methods of quantifying change detection accuracy such as that proposed by Tang et al. [283].

A possible improvement to the simulated datasets used in this study would be to include more realistic year-to-year and seasonal variations. For example, seasonal cycles are likely to have yearly variations in amplitude, length, and shape due to fluctuating weather conditions or variations in yield. Noise was also distributed evenly throughout the year, whereas in reality noise caused by factors such as cloud contamination tend to be clustered around certain times of year.

In addition to creating more realistic simulations, there is also potential to use simulated data to better explore the limitations of individual methods. Many methods have multiple parameters that must be set by the user. While expert knowledge of the study area can be used to decide these values, simulations can support this knowl-

edge by allowing the user to test how different parameters might cause an algorithm to behave differently under different scenarios. Simulations can also be parameterised to better reflect specific study areas or vegetation types [75, 296, 297]. This type of customisation could also be improved using the suggestions given above, for example, by estimating how much year-to-year seasonal variation should be expected in vegetation by using climate data.

The simulations used in this study provide a starting point for future studies and have been made available for download [16].

3.4 Conclusions

A novel means of robustly evaluating and comparing change detection techniques using simulated time series data is presented in this study. This firstly allows for each method to be evaluated against a wide range of change scenarios, including those where data are noisy or incomplete. Secondly, this process allows for comparison between methods based on temporal accuracy, likelihood of detecting false changes, and RMSE number of breaks. The insights gained can be used to provide recommendations for users as to which method might be most appropriate for their application. However, due to the limitations of this study, it is important to emphasise that further investigation and optimisation should be carried out to ensure the efficacy of any method when applied to a specific use-case. In particular, the selection of input parameters such as the order of the seasonal component, the number of breakpoints to detect, the length of the history period, and the value of λ for CCDC and EWMACD will have a substantial impact on the results achieved and in many cases default values will not be the most appropriate.

Recommendations

- For smaller magnitude changes such as partial forest harvesting within pixels and for detecting changes in land cover condition (e.g., due to decreasing yield or recovery after fire), EWMACD is likely to be the most effective due to its ability to detect a wide variety of change magnitudes and low false detection rate.

- For studies which aim to robustly detect complete changes in land cover class (e.g., change from forestry to cropland), CCDC with a fixed λ is recommended. CCDC performed well at detecting larger magnitude changes and tended to ignore or underestimate smaller magnitude changes and seasonal changes. Using fixed λ greatly increases algorithm runtime, although λ should be chosen carefully in order to maximise or minimise change detection as appropriate.

- The detection of seasonal changes is a field in itself and software packages such as TIMESAT [142, 143] can aid in more detailed reconstruction of seasonal curves. However, of the methods investigated here, it was found both EWMACD and BFAST Monitor capable of detecting at least high magnitude seasonal change, such as a change in the number of seasonal peaks present (indicating a change in cropping practices) or a substantial increase in seasonal amplitude (indicating, e.g., a change in yield). Of the two, EWMACD is recommended due to its lower likelihood of detecting false change.

- If data are known to be noisy, e.g. with many small clouds or cloud shadows present which are difficult to screen out, either EWMACD or BFAST could be suitable. EWMACD was able to find more correct breaks in time series regardless of the level of noise, whereas BFAST was the most consistent method across noise levels for all metrics. However, given the poor performance of

BFAST on the seasonal change sets, its use is only recommended here for finding abrupt changes.

- For datasets with high levels of missing observations such as those from areas of the world with high year-round cloud or snow cover, CCDC is recommended. CCDC gave very consistent performance across missing data levels, probably because it is designed to look for land cover class changes and is less likely to be influenced by single outliers. The adaptive Lasso regression method should also help to correctly estimate seasonal parameters if data are missing.

- As computing power increases, change detection techniques can be applied across larger and larger datasets. Most of the methods discussed here are now available on Google Earth Engine [104]. Initiatives such as the Open Data Cube show the potential of continental scale analysis [165]. However, pixel-level change detection is still computationally expensive. Based on its good overall performance and fast execution time across multiple change types, EWMACD shows potential for large scale analysis.

3.4.1 Applicability to mangrove monitoring

The above results provide insight into the best methodology to apply to the specific use-case of large scale mangrove monitoring. As discussed in Chapter 2, mangrove ecosystems across the world are highly variable and dynamic, and experience a variety of change pressures. These pressures are represented in this Chapter within the various change sets. Establishing mangrove stands, or stands recovering after a major event such as a cyclone are likely to exhibit positive NDVI trends. Mangroves affected by longer term environmental effects such as oil spills or salinity changes will exhibit decreasing NDVI trends. These trends could be gradual, for example due to changes in species composition leading to a denser vegetation understory, or be caused or interrupted by sudden weather events such as flooding or strong winds.

The latter case is represented by the break/trend set. Severe cyclone impacts from which mangroves recover, or deforestation and re-establishment practices as seen in the Matang Mangrove Forest Reserve, are represented by the *Extreme* break set. This set could also represent a loss of mangroves as impacts wash away soil and prevent re-colonization. A *Moderate* break in a mangrove NDVI time series might be caused by a less severe storm event or one from which mangroves struggle to recover (represented by simulations with a medium break followed by a weak trend), for example due to soil contamination. Both *Extreme* and *Moderate* breaks could lead to land cover class change to soil or water if the stand fails to recover. *Subtle* breaks could indicate partially changed pixels (e.g. small-scale felling) or changes that affect only some mangrove species present (e.g. because of a change in salinity). In this case the break would be unlikely to lead to a change in land cover class.

Although they are evergreen trees, mangroves do exhibit seasonality, with greenness and litterfall being closely related to environmental variables such as temperature and rainfall [225]. Changes in SOS and peak productivity (seasonal amplitude) are relevant to any vegetative land cover type which could be affected by climate change. In mangroves, such seasonal changes may be hard to detect and indicate long-term shifts in phenology due to changes in sea/water temperature and sea salinity.

Taking into account the types of change present in mangrove ecosystems, while EWMACD gave good performance overall, CCDC was ultimately chosen as the most appropriate method for further study. A major reason for this is that CCDC gave the most consistent performance across missing data levels, and a lack of available observations due to high cloud cover is one of the main potential limitations of time series modelling in tropical and sub-tropical areas. CCDC provides the highest likelihood of achieving a good result even for study sites with substantial and persistent cloud cover. CCDC is also more robust to smaller magnitude changes, which could indicate partial change or ephemeral change rather than land cover change, and had the low-

est overall rate of false break detection of all the methods studied. In addition, for dynamic mangrove ecosystems which are constantly in flux, EWMACD could struggle to establish a period of general stability from which change could be detected. EWMACD also has a higher requirement for tuning to specific scenarios, as the definition of change is strongly influenced by the chosen control limits and time series weighting [35]. In contrast, CCDC is designed to detect general land cover change with little tuning [340]. CCDC was therefore considered to be the best choice for large-scale monitoring of change in mangrove forests, where robust detection of class changes across a range of different study areas is more useful than a more site-specific approach.

Chapter 4

Using Continuous Change Detection and Classification of Landsat Data to Investigate Long Term Mangrove Dynamics in the Sundarbans Region

4.1 Introduction

Mangroves are salt-tolerant trees which occupy the intertidal zone. Globally mangrove forests cover an area of nearly 140,000 km^2 [39, 94] in over 120 countries and territories [160]. These forests play an important role as carbon sinks [158] in addition to providing many ecosystem goods and services [236]. Occupying a narrow ecological niche, mangroves are particularly vulnerable to climate change effects such as sea level rise, changing ocean currents, and increasing temperatures, which lead to greater erosion, increased salinity, and reduced sediment deposition [306]. Globally, many mangrove ecosystems remain threatened by anthropogenic activity [74]. Understanding the dynamic nature of mangrove ecosystems is vital for both preservation and for utilization of theses ecosystems as climate change markers.

While many studies exist which aim to monitor long-term change in mangrove ecosystems (e.g., [8, 97, 98, 172, 288], the focus has remained on single-image classification techniques, typically at yearly resolution or less. In recent years land cover

monitoring has seen a move away from such approaches and towards methods which make use of all available observations [338]. Method such as Continuous Change Detection and Classification (CCDC) [340, 342] use a per-pixel model fitting approach to capture the seasonal dynamics of land cover types in addition to their inter-year greening and browning trends. One advantage of model fitting is that it reduces reliance on individual observations, instead attempting to capture the broad phenological cycle of the underlying vegetation. It also reduces the need for whole cloud-free images, instead relying on each individual pixel. This is particularly useful for monitoring mangrove ecosystems, which typically exist in tropical and sub-tropical regions where high cloud cover reduces observation frequency [160]. Changes in land cover can then be deduced by finding where in time the fitted season-trend model breaks down. In addition, a classification of the different models can be made based on their various parameters. Unlike a single-image classification approach, this type of classification takes into account the temporal signature of the vegetation rather than spectral values at a single point in time, potentially leading to higher classification accuracy.

Spanning parts of both Bangladesh and India, the Sundarbans is the largest contiguous mangrove forest in the world [55, 61, 130]. It is an area of high biodiversity [209] which provides many ecosystem goods and services, such as timber for building and fuel [90]; materials for medicines [209]; stabilization of the coastline and protection against extreme events such as cyclones and tsunamis [188]; maintaining water quality [99]; and habitat for a variety of endangered species. More than 5 million people depend on the Sundarbans mangroves for resources [209]. While much of the Sundarbans has been designated a United Nations Educational, Scientific and Cultural Organization (UNESCO) World Heritage Site [208], it remains threatened by anthropogenic activities such as over-exploitation, illegal logging, pollution, and expanding industries such as shrimp farming. In addition, rising sea levels are increasing

salinity in the delta, with adverse effects on the dominant mangrove species, *Heritiera fomes* [24, 91]. There is also evidence that the frequency of tropical cyclones making landfall in Bangladesh is increasing [134]. Damage from such cyclones, in combination with rising sea levels, means that the protective capacity of the Sundarbans could be reduced in future years. High importance must therefore be placed on understanding the potential long-term impact of cyclones on the Sundarbans mangroves in addition to tracking mangrove extent.

Given the importance of the Sundarbans, both in terms of providing resources and protecting coastal communities, adequate monitoring of mangrove abundance and health is vital. However, as for many mangrove forests, field studies are expensive and time consuming due to the size of the area covered and difficulty of access [91, 97]. The Sundarbans is also an extremely dynamic region which experiences rapid changes due to tides and flooding from tidal surges and seasonal rains, in addition to long-term erosion and accretion of land [97]. These dynamics are difficult to capture on the ground, where a simultaneous snapshot of the entire area is not possible. Remote sensing offers a solution to this. Datasets such as the Landsat archive can provide data going back over multiple decades at a high enough spatiotemporal resolution to identify and track localized changes, both current and historical. In addition, as the price of computing infrastructure continues to fall, the processing of large areas such as the Sundarbans is becoming faster and more viable.

Many studies have used satellite imagery to monitor dynamics in the Sundarbans in recent years. These have covered a wide range of topics including overall mangrove extent [97, 98], land cover and species level change [60, 91, 99], mangrove phenology [11], and coastline change [243]. However, these studies are limited in that they only use one or two images per year [60, 91, 98, 99], do not cover the whole extent of mangroves in the Sundarbans region [99, 243], or use data at lower spatial resolution than Landsat [11]. While several studies attempt to quantify damage and recovery

from cyclones (e.g., [6, 28, 55]), the focus on a single image or comparison between very few before/after images is limiting in that it does not take into account the wider long-term variability of the Sundarbans area.

Mangroves typically exists in tropical and sub-tropical regions which are often covered by clouds. This creates difficulty in obtaining cloud-free images from the same time of year which can be compared. For this reason, Cornforth *et al.* [55] used Synthetic Aperture Radar (SAR) data to study damage and recovery from a major cyclone in the Sundarbans. However, there are disadvantages to the use of SAR imagery. Firstly, no single SAR dataset offers the continuous decades-long coverage of Landsat, making long-term monitoring difficult. Secondly, while SAR data can provide information about tree canopy structure which is absent from optical data, there is evidence that optical data can provide better discrimination between mangroves and other similar tropical vegetation than SAR data [39]. In central areas of mangrove forest where vegetation is overwhelmingly likely to be mangroves, discrimination between land, water, and vegetation is adequate. However, for detecting smaller clusters of mangroves over wider areas a more comprehensive classification is needed.

There is clearly scope for a more in-depth study which makes full use of the temporal record for mangrove ecosystem monitoring. In this study, mangrove extent and greening and browning trends in the entire Sundarbans region are examined over thirty years using the Landsat data archive and CCDC. Damage extent and recovery from a major cyclone, Sidr, is also investigated. By taking advantage of all available observations at a per-pixel level, this study aims to provide a more comprehensive and accurate view of mangrove dynamics in the Sundarbans. In doing so, the utility of using a dense time series method for both mangrove classification and monitoring over a large area is demonstrated.

4.2 Materials and Methods

4.2.1 Study Area

The study area covers the region from 20°8′ to 24°0′N and 87°4′ to 93°0′E (Figure 4.1). This covers the entirety of the Sundarbans, approximately two-thirds of which is located in Bangladesh and the remaining third in India [90]. Located in the Bay of Bengal at the confluence of several major river systems (the Ganges, Brahmaputra and Meghna), the Sundarbans is a low-lying area of dense mangrove forest intersected by a complex network of river channels, islands, and mudflats. Elevation is between 0.9 and 2.1 m above Mean Sea Level (MSL) [129]. The climate is tropical, with the monsoon season extending from May to October [91]. The area is periodically hit by cyclones [134], many of which cause substantial losses of both lives and property [226].

The Sundarbans is an area of high biodiversity, home to many rare, threatened, or endangered species [103]. Three main mangrove species dominate: *Heritiera fomes*, which can only tolerate low water salinity; *Excoecaria agallocha*, which can grow in moderately saline water; and *Ceriops decandra*, which can tolerate high salinity [11]. These species have been estimated to make up 33.4%, 30.2%, and 32.4% of mangroves in the Sundarbans, respectively [91]. While there is evidence that numbers of *Heritiera fomes* and *Excoecaria agallocha* trees are declining, numbers of *Ceriops decandra* may be increasing [91].

4.2.2 Data and Pre-Processing

All United States Geological Survey (USGS) Collection 1 Tier 1 Landsat 4-5 TM, Landsat 7 Enhanced Thematic Mapper Plus (ETM+) and Landsat 8 Operational Land Imager (OLI) data covering the Sundarbans national park region for the period from January 1988 to June 2018 was downloaded from the Google public repository. Collection 1 Tier 1 provides high quality data which has been georegistered

and inter-calibrated across the Landsat instruments, and is considered suitable for time series analysis [291]. The Sundarbans is covered by six Landsat scenes: WRS-2 (World Reference System 2) Row 45/Path 137, Row 44/Path 137, Row 44/Path 138, Row 45/Path 138, Row 45/Path 136, and Row 44/Path 136 (Figure 4.1). The downloaded images were atmospherically and radiometrically corrected converted to analysis ready surface reflectance data using the Atmospheric and Radiometric Correction of Satellite Imagery (ARCSI) Python package [38]. As part of this process cloud masks were also created using the Function of mask (Fmask) algorithm [341].

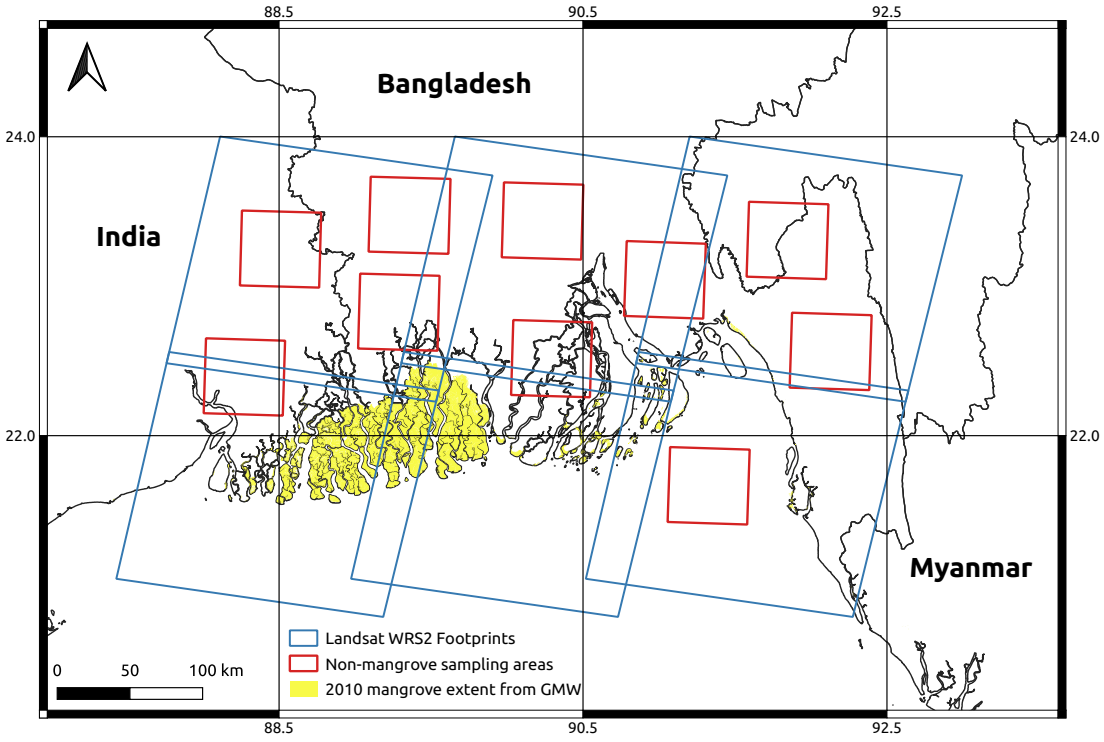


Figure 4.1: .
Overview of the study area. The Sundarbans region is covered by six Landsat scenes (blue squares). Red squares indicate the areas from which non-mangrove training data was selected, where each square has an area of approximately 3000 km^2. Mangrove training data was taken from the Global Mangrove Watch (GMW) 2010 classification.

The data were then ingested into a data cube on Super Computing Wales (SCW). The Open Data Cube (ODC) is an open source program which uses a database along

with a Python interface to simplify the processing and organization of geospatial data [18]. Data is tiled and spatially aligned to allow for easy analysis of per-pixel time series. Processing can also be parallelized to facilitate analysis of large scale and even continental scale datasets [165]. As part of the process of data cube ingestion, the data were divided into 105 × 105 km tiles (roughly 1°by 1°) and re-projected to the Bangladesh Transverse Mercator projection (EPSG 3106). It was then stored as Network Common Data Form (NetCDF) files on SCW ready for analysis. Based on work by Zhu *et al.* [342] only the red, green, near-infrared (NIR), and shortwave infrared (SWIR) bands were used for analysis.

4.2.3 Classification of Mangroves Using CCDC

4.2.3.1 The CCDC Algorithm

Once data had been pre-processed, the Continuous Change Detection and Classification (CCDC) method was applied to the Sundarbans area to provide data on the timing and magnitude of class changes (e.g., mangrove to non-mangrove) in addition to details on trend over time (Figure 4.2). CCDC is designed to work with multi-band Landsat data and focuses on changes in land cover class whereby breaks in the time series are identified and each segment (period between breaks) is classified independently [340]. Breaks are found by fitting a linear model to a stable history period of up to two years. New observations are then added and their residuals compared to the RMSE (Root Mean Square Error) of the history period. If the error of a new observation is too high, it is flagged as a potential change. Once six such observations are flagged in a row a change is identified and a new stable model is initialized. Six observations are used to provide a high confidence of change having occurred [342]. Tropical forests present a challenge for change detection analysis because the prevalence of clouds reduces data frequency. While CCDC has been applied to tropical forest monitoring [283], it has not previously been used for analysis of long-term

trends in tropical vegetation or for classifying mangrove forest. However, given the length of the time series used some effects of missing data should be mitigated because observations from different years are unlikely to fall on the same day of year.

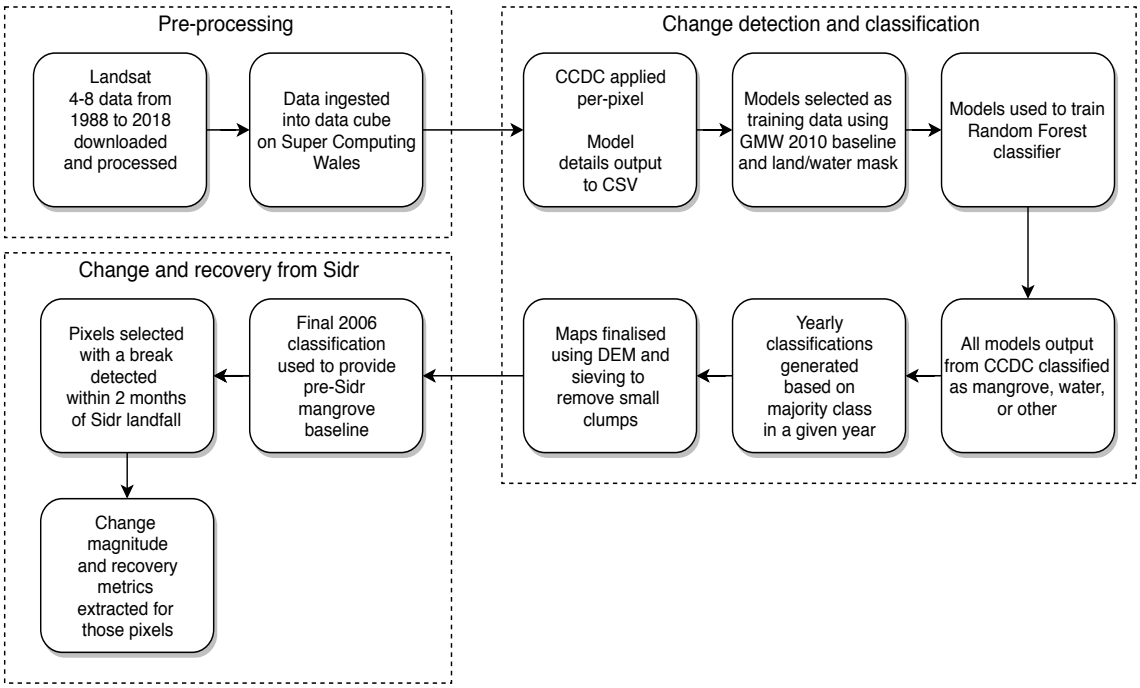


Figure 4.2: Overview of processing steps.

CCDC is a computationally expensive method [340]. The area processed is approximately 600 × 380 km in size constituting nearly 250,000,000 pixels. To process the area in a reasonable time frame, the SCW facility was used. This provided access to up to 600 cores simultaneously allowing for parallel processing of pixels. Running CCDC over the whole area took around two weeks, however, 600 cores were not continually available during this period.

The version of CCDC used was that proposed in [342] and was implemented in Python 3.6.8. This version fits a season-trend model of the form given in Equation (4.1), where $\hat{\rho}(i, x)$ is the predicted value for the ith Landsat band at Julian date x, $a_{0,i}$ is the coefficient for the mean of the ith Landsat band, $a_{1,i}$ and $b_{1,i}$ are coefficients representing intra-annual change, and $c_{1,i}x$ is the coefficient representing the inter-annual change, or trend [340]. $T = 365.25$ (the number of days in a year). CCDC is

adaptive, including up to three harmonic (sine/cosine) terms in the model depending on available data quantity [342]. The model is fitted separately to the data for each Landsat band using Lasso regression. Lasso regression minimizes overfitting through regularization by limiting the magnitude of the model coefficients [342]. The degree of regularization depends on the value of the parameter λ where $0 < \lambda < \infty$. If the value of λ is too high, the model will underfit and fail to capture the seasonal nature of the data. If λ is too low, the model could overfit, resulting in a higher rate of false changes being detected. The value of λ was decided through manual interpretation of models fitted to a sample of mangrove time series where $\lambda = 0.01$, 0.1, 1, 5, and 20. A value of 1 was chosen in this case because it resulted in models which adequately captured the seasonal cycle of the data while minimizing overfitting.

$$\hat{\rho}(i,x) = a_{0,i} + a_{1,i}cos\left(\frac{2\pi}{T}x\right) + b_{1,i}sin\left(\frac{2\pi}{T}x\right) + c_{1,i}x \tag{4.1}$$

4.2.3.2 Generation of Yearly Classification Maps

Classification maps identifying pixels as either mangrove, other terrestrial, or water were generated for each year from 1988 to 2017 (inclusive). A classification map was not created for 2018 because data had only been processed up to June 2018. A full year of data was, therefore, not available.

Model Classification In addition to dates of change for each pixel through time, the algorithm outputs were per-band model coefficients ($a_1, b_1 ... a_3, b_3$), RMSE, and an overall value for each model calculated using the slope and intercept [340]. Given an input of five image bands and a third-order harmonic model, this resulted in a set of 45 variables for each model. These variables were used to classify each model as mangrove, water, or other terrestrial using the method provided by Zhu *et al.* [340]. The assumption of this method is that different land cover types have different

seasonal cycles, resulting in different model coefficients and different model fits. A classifier can therefore be trained to identify different land cover types based on the fitted season-trend model.

First training data was obtained using the Global Mangrove Watch (GMW) Version 2.0 dataset. The GMW provides a highly accurate global mangrove baseline for 2010 [39]. A sample of 50,000 pixels was taken randomly from within the GMW 2010 mangrove area. To obtain a non-mangrove sample, ten polygons of approximately 3000 km^2 in area were then chosen from around the Sundarbans region to maximize the distribution of the training data (Figure 4.1). Data were not taken from the entire non-GMW region due to the risk of including mangrove pixels not identified by the GMW baseline. A sample of 50,000 pixels was then randomly selected from each polygon to obtain a total sample of 500,000 pixels. The sample size was large enough to ensure that most dominant land cover classes were represented. Once the training pixels had been selected, the results of the change detection step were used to identify whether those pixels had fitted models suitable for use as training data.

Each model generated by CCDC covers a specific period of time for a specific pixel. Therefore, in order to classify models based on the 2010 dataset, only models which covered the whole of 2010 were chosen as training data because a break detected in 2010 could signal a change in land cover type. There would then be no way of knowing which model aligned with the mangrove/non-mangrove classification provided by GMW. If a model covered a period starting before 2010 and ending after 2010, that model likely represented a mangrove/non-mangrove signal as defined by the 2010 data. The non-mangrove training models were further divided into land and water models using a land/water mask for 2010. This gave final final training datasets of 49,923 mangrove models, 87,736 water models, and 403,088 other terrestrial models.

Once a set of stable models was identified, the model parameters were used to train a Random Forest (RF) classifier implemented using the scikit-learn Python library

[227]. The same model parameters suggested by Zhu *et al.* [340] were used, however in almost all cases CCDC fitted a third-order harmonic model resulting in a set of nine variables per band or 45 total variables used for classification per model. In the rare cases where a less complex model was fitted the extra coefficients were set to zero to maintain compatibility with the classifier. To find the best hyperparameters for the classifier, a random grid search with cross validation was used to narrow down the potential range of values. A second grid search was then performed to produce the final set of hyperparameters. Once the classifier was trained, all models produced as output from running the CCDC algorithm per-pixel were classified.

Yearly Maps The output of the CCDC algorithm resulted in a set of classified models per-pixel, each with a start and end date. Some models covered only a few years whereas others covered the entire time period. To generate yearly classification maps it was necessary to summarize these models into a yearly class for each pixel. To achieve this the majority class within a given year was used. If there was no majority, the pixel was given a value of 0 (not enough data) for that year. Mangrove pixels were given a value of 1, other terrestrial pixels a value of 2, and water pixels a value of 3. The resulting yearly maps were then processed using the Shuttle Radar Topography Mission (SRTM) global 1 arc second product to remove any mangrove pixels above 30 m. Mangroves only grow at low elevations in tidal and intertidal zones [98], so any pixels above this height were highly likely to be miss-classifications. Finally, any clumps of land cover less than 10 pixels (0.9 ha) in size were removed and replaced with the value of the largest neighboring class using the Geospatial Data Abstraction Library (GDAL) [83], in line with recommendations by Bunting *et al.* for reducing error associated with small-scale features [39]. Yearly mangrove extent in km^2 was then calculated based on the number of pixels in each yearly map assigned to class 1.

Validation It is important in any classification that accuracy is independently assessed by a validation dataset. Due to the spatiotemporal extent of the data, validation using field data or a higher resolution dataset was not possible. No other dataset is available which covers the same time period as Landsat at similar or higher spatial resolution. Following the same methodology as the GMW [39], a validation set was therefore generated by taking a randomly selected set of 47,000 pixels sampled across both space and time and interpreting them manually using Landsat data. Randomly sampling over the whole Sundarbans area could result in a biased validation sample as mangroves make up a minority of the study site. Therefore, sampling was carried out using a set of pre-defined polygons selected to ensure good coverage of the mangrove forest itself in addition to surrounding water and vegetation. Reference to up to date high resolution Google Earth imagery was made and the Landsat imagery was displayed using the NIR, SWIR1 and Red colour composite, in which the mangroves are spectrally distinct (see Figures 4.4 and 4.5), to aid identification of mangroves from other terrestrial land covers. Each validation pixel was assigned a class (mangrove, water, or other terrestrial vegetation) through expert human interpretation. Where there was uncertainty about the true classification for a pixel, the highest likelihood class was assigned based on spatial and temporal context. These classes were then compared to the final yearly classification generated by CCDC.

The CCDC classification resulted in some pixels for which no overall land cover class could be determined (i.e., there was no majority class in the year) so the final number of validation points used was 45,440. User's and producer's accuracy were calculated for each of the three classes along with overall accuracy. Quantity disagreement and allocation disagreement were calculated as described by Pontius and Millones [237].

4.2.4 Long Term Vegetation Trends in the Sundarbans

As well as determining yearly classifications, the models created by the CCDC process were used to generate a mangrove trend map covering January 1988 to June 2018. The purpose of this was to summarize greening and browning trends for the Sundarbans mangroves both spatially and temporally. Using the yearly classification maps, all pixels which had been classified as mangrove at any point between 1988 and 2017 were identified. For each pixel, the overall Normalized Difference Vegetation Index (NDVI) trend was then calculated using the method described by Zhu *et al.* [345] over all models for that pixel. Briefly, this method uses the intercept and slope of the fitted seasonal models for the red and NIR bands to calculate an overall NDVI value for the start and end date of each model. Overall trend is then calculated by adding up the within-model and between-model differences.

4.2.5 Investigation of Dynamics Around Cyclone Sidr

In addition to investigating the accuracy of CCDC for mangrove classification in the Sundarbans, a specific use case was also desirable to illustrate the potential for more in-depth investigation of vegetation dynamics. Given the computational complexity of CCDC, its usefulness for many applications could be increased if multiple analyses can be carried out using the same results. In this case disturbance and recovery dynamics around Cyclone Sidr were investigated as the final stage of analysis (Figure 4.2).

Cyclone Sidr was a category 5 storm which made landfall in Bangladesh on the 15th of November 2007. It caused a 3 to 4 m tidal surge along with wind speeds up to 220 km/h [6]. Sidr was chosen as a use case for several reasons. Firstly, the cyclone hit the Sundarbans area directly [226]. Secondly, Sidr hit in 2007, meaning that data were available for multiple years both before and after it made landfall. Thirdly, estimates of damage done by Sidr around the time of impact range widely

from around 22% to 45% of mangroves [28, 55] and no research was found into the long-term effects of cyclone Sidr beyond 2010.

In order to investigate the possible long-term effects of Sidr, all pixels classified by CCDC as mangrove in 2006 were selected. From this set all pixels were selected for which CCDC recorded a break in the two months following Sidr. Two months was chosen to provide a reasonable level of certainty that the change was attributable to Sidr rather than other causes, given the lack of available validation data. The magnitude of the reported break was recorded along with the classification of the first post-Sidr model and the number of days between the end of the pre-Sidr model and the start of the first post-Sidr model. The trend method outlined in Section 4.2.4 was then used to estimate an overall NDVI value for the end of the final pre-Sidr model, i.e., a greenness value for the pixel before the cyclone hit. The trend for each subsequent model was then tracked to find the first year in which the overall NDVI value reached or exceeded the final pre-Sidr value. This year was then recorded as the year of recovery from Sidr. If the pixel had not recovered by the end of the available data in mid-2018, it was recorded as Not Yet Recovered (NYR).

4.3 Results

4.3.1 Classification of Mangroves Using CCDC

The classification achieved an overall accuracy of 94.5% with a 99% confidence of being between 94.2% and 94.8% (Table 4.1). Quantity disagreement was 0.02 and allocation disagreement was 0.04. Mangroves were most frequently confused with other terrestrial pixels. Other terrestrial pixels were most often confused with water.

Based on the yearly classification maps, the year with the highest mangrove extent was 2004 with mangroves covering approximately 6672.6 km^2. The year with the lowest mangrove extent was 1988 when mangroves covered approximately 6368.3 km^2. The lower values for 1988 and 1989 (Figure 4.3) are probably because Bangladesh

experienced both catastrophic flooding and a severe cyclone in 1988. These events caused rapid changes in land cover (i.e., water inundation) and destruction to vegetation making it difficult to fit a stable model. Mangrove extent also increased substantially between 1989 and 1990 (Figure 4.3), possibly because the forest was still recovering. Excluding 1988 and 1989, the year of lowest extent was 2007 (6576.0 km^2). Mangrove extent increased by 15.1 km^2 between 1996 and 2006 (0.23%), the ten years before Sidr, and decreased by 40.9 km^2 between 2008 and 2017 (0.61%). However, the majority of this decrease was between 2013 and 2017, with extent only decreasing by 0.03% between 2008 and 2013. In the last 20 years, between 1997 and 2017, mangrove extent reduced by 38.4 km^2 (0.58%). All of these figures are lower than the classification uncertainty and should therefore be interpreted cautiously. A full table of yearly extent change can be found in Table A.1.

Table 4.1: Results of spatiotemporal accuracy assessment for CCDC.

		Reference				
		Mangrove	**Water**	**Other**	**Total**	**User's (%)**
Classifier	Mangrove	11,835	531	219	12,585	94.0
	Water	174	22,422	367	22,963	97.6
	Other	684	609	8599	9892	86.9
	Total	12,693	23,562	9185	45,440	
	Producer's (%)	93.2	95.2	93.6		**94.5**

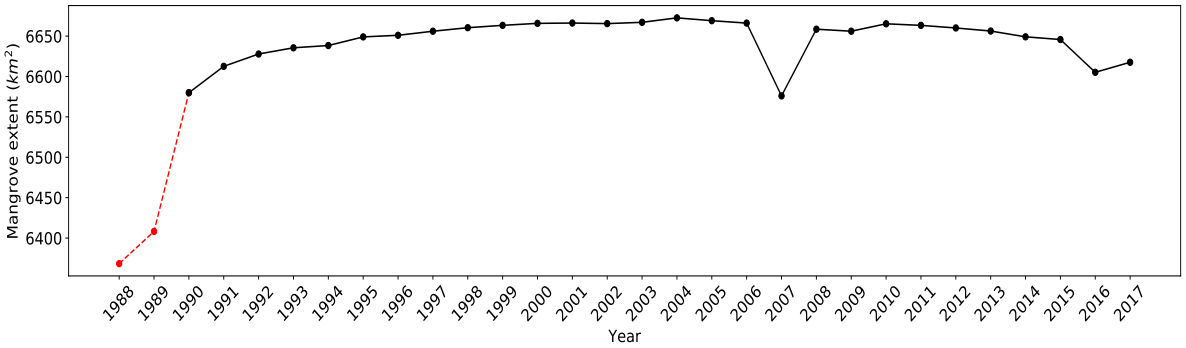


Figure 4.3: Area covered by mangroves over time, based on the final yearly classifications. Extent for 1988 and 1989 is represented by a dotted red line because the classification of mangroves in these years was considered less reliable due to a combination of the extreme weather experienced by Bangladesh in 1988 and spin-up effects of the modelling approach.

For 2010 there was substantial overlap between CCDC and the GMW. Only 1% of mangrove area captured by the GMW was not captured by CCDC. However, mangrove extent derived using CCDC for 2010 (6665.2 km^2) was much higher than estimated by the GMW (6016.4 km^2). 10.7% of the mangrove area captured by CCDC was not classified as mangroves by the GMW, indicating that the CCDC approach was able to capture areas of mangrove which were missed by the GMW (Figure 4.4). The CCDC classification was also able to capture long-term changes in mangrove extent caused by land erosion and accretion (Figure 4.5). A specific example is shown in Figure 4.6D of a pixel which developed from water to mangrove and was robust enough to recover from a major disturbance event.

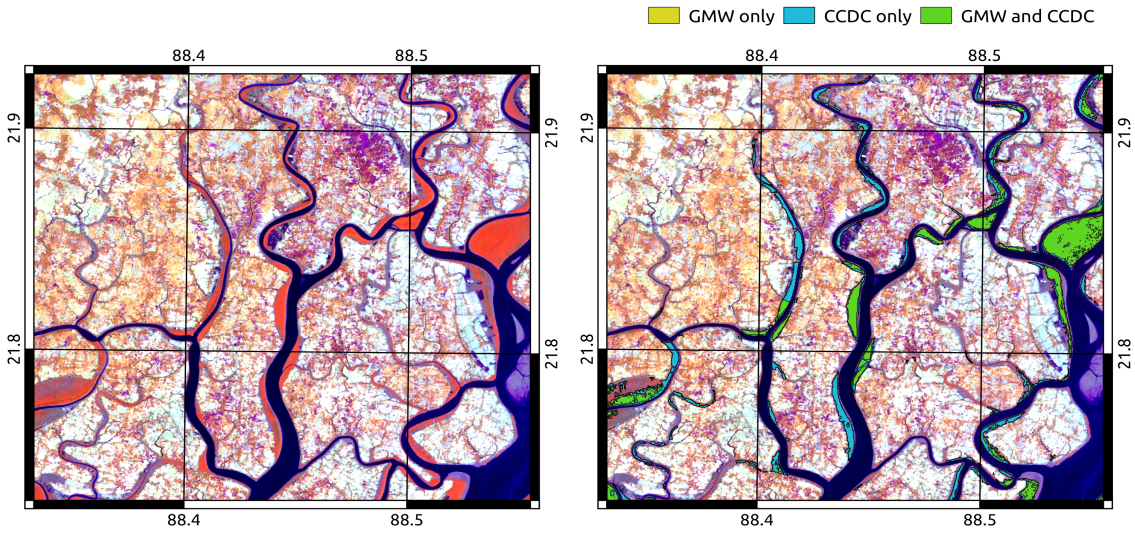


Figure 4.4: Left: A false colour Landsat 5 image from February 2010 where R = NIR, G = SWIR1, and B = Red. Right: Comparison of CCDC classification vs. GMW classification for 2010 showing additional mangroves captured by CCDC.

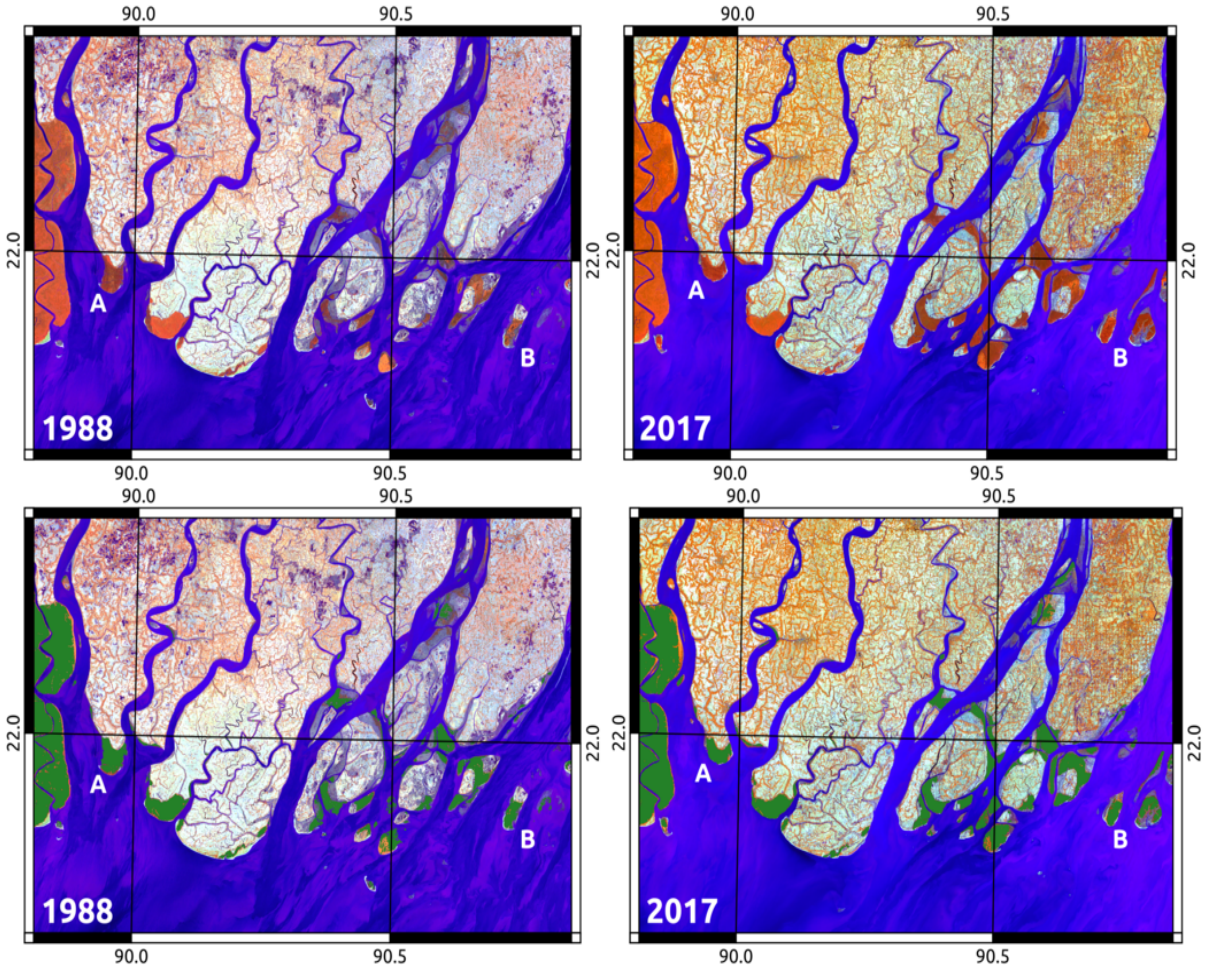


Figure 4.5: Examples of loss and gain of mangroves between the 1988 and 2017 classification maps. The top row shows a false colour RGB image where R = NIR, G = SWIR1, and B = Red. The bottom row shows pixels classified as mangrove by CCDC in green. Location A shows a loss of mangroves while location B shows formation of new land with mangroves establishing.

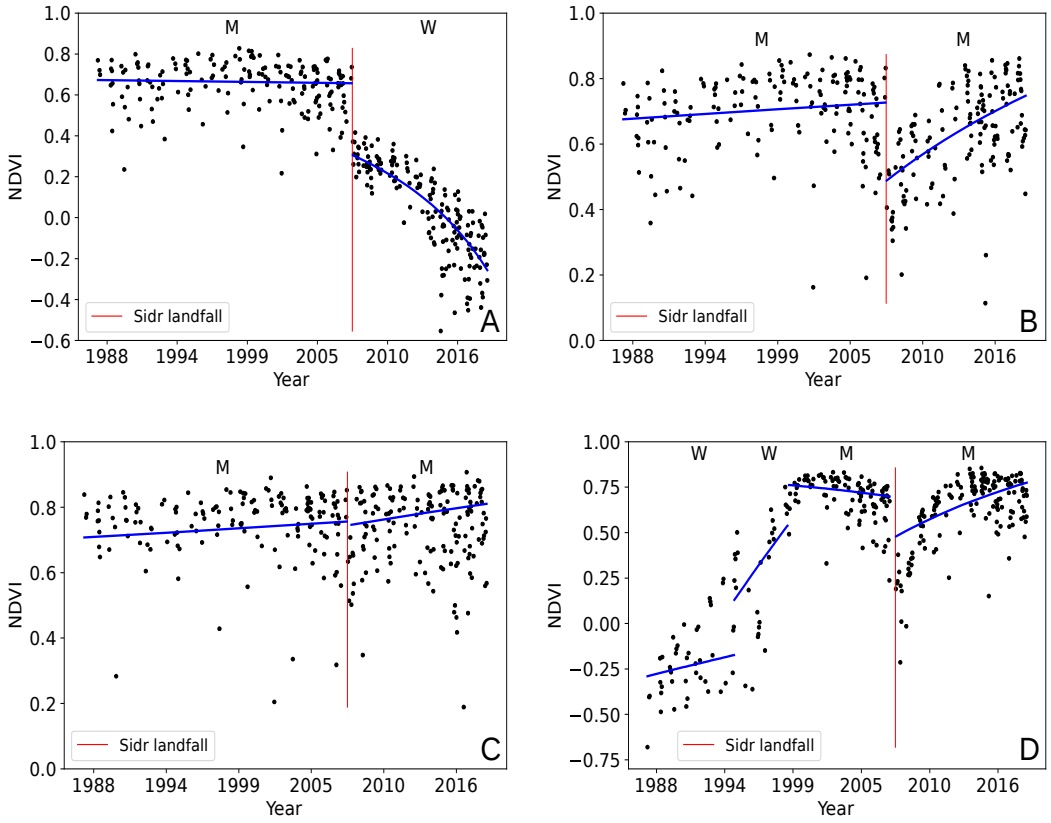


Figure 4.6: Example NDVI trajectories for four pixels where CCDC recorded a break in the 2 months after Sidr made landfall. (**A**) Pixel that did not recover from Sidr and was classified as water once a new model had been initialized; (**B**) Pixel that was significantly damaged by Sidr but recovered and was still classified as mangrove once a new model had been initialized; (**C**) Pixel showing minor damage from Sidr; (**D**) Pixel showing a possible area of new land formation where an establishing trend in NDVI can be seen. Once established the pixel remains classified as mangrove even after a high magnitude disturbance. M = Classified as mangroves, W = classified as water.

4.3.2 Long Term Vegetation Trends in the Sundarbans

Approximately 0.002% of pixels reported an unrealistic overall trend (trend < -2 or trend > 2) and were removed from the analysis. These trends were probably caused by outliers (e.g., very high or negative values) in the red and NIR bands. Of all remaining pixels classified as mangrove at any point between 1988 and 2017 (6865.6 km^2), 73.5% experienced an overall positive trend in NDVI (5044.1 km^2) and 26.5% experienced an overall negative trend (1821.3 km^2). The median positive trend value was 0.06 with 24.6% of positive trend pixels (18.1% of all pixels) exhibiting a trend

>= 0.1 and 4.5% of positive trend pixels (3.3% of all pixels) exhibiting a trend >= 0.5. The median negative trend value was −0.03 with 19.5% of negative trend pixels (5.2% of all pixels) exhibiting a trend <= −0.1 and 7.8% of negative trend pixels (2.1% of all pixels) exhibiting a trend <= −0.5. These figures indicate that more greening than browning is occurring across the Sundarbans mangroves. Spatially, stronger greening and browning trends occur around the coastline with a more generalized area of browning also apparent to the north (Figure 4.7).

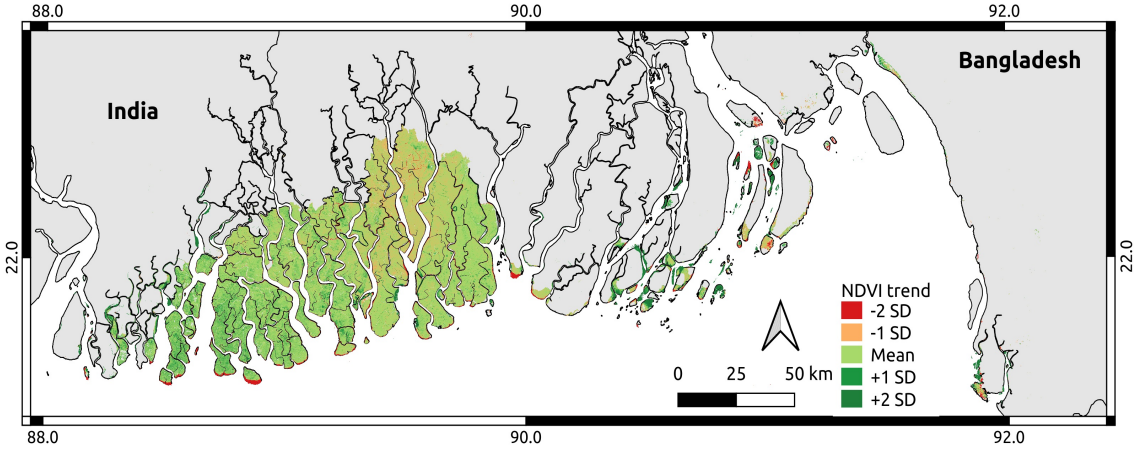


Figure 4.7: Overall NDVI trend from January 1988 to June 2018 for all pixels classified as mangrove at some point (union of all classifications). Stretched to $\pm 2\sigma$ to better show trend distribution.

4.3.3 Investigation of Dynamics Around Cyclone Sidr

An estimated 10.7% of mangroves were damaged by Cyclone Sidr based on change from the 2006 classification. Damaged areas were clustered around the coastline and to the west of the cyclone's path (Figure 4.8). The median NDVI break magnitude was -0.1 with an interquartile range of 0.05. Break magnitude cannot be directly related to the amount of damage caused on the ground. However, 51.9% of pixels registered a break magnitude >= −0.1, indicating that the majority of disturbances were small (Figure 4.6C). Median days disturbed (days before CCDC was able to fit a new stable model) was 80, indicating that land cover often took more than two months to recover to a level stable enough for a new model to be fitted. Minimum days disturbed was

one. This was possible because a small amount of overlap between Landsat 5 and Landsat 7 scenes in January 2008 meant that some pixels had observations on two consecutive dates. Maximum days disturbed was 1040. 88.6% of pixels had a new model fitted within 180 days.

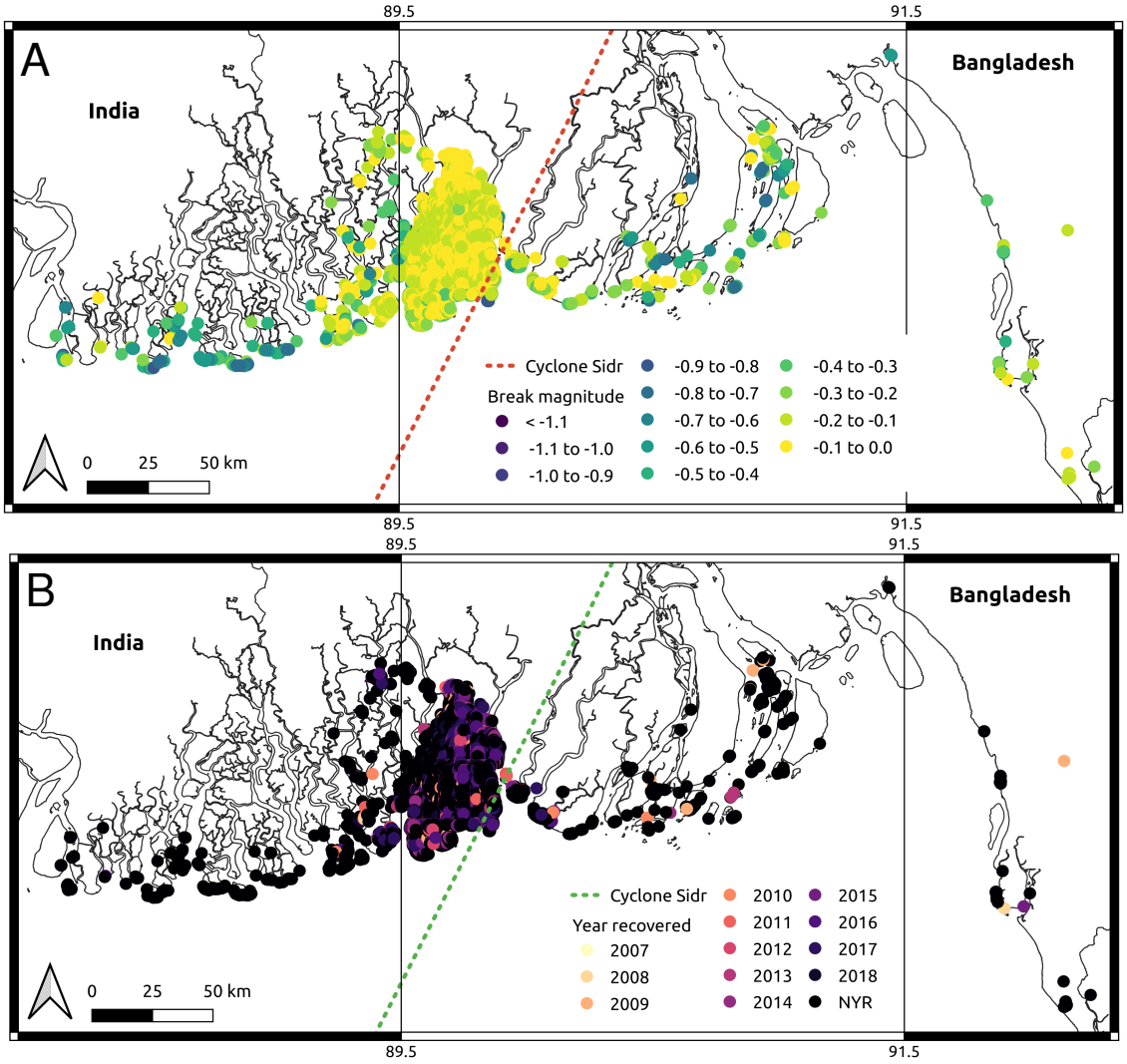


Figure 4.8: Maps showing (A) distribution of NDVI break magnitude recorded by CCDC and (B) distribution of year recovered from Cyclone Sidr in terms of NDVI trend. NYR = Not Yet Recovered.

95.7% of damaged pixels were classified as mangrove once a new model had been fitted (684.0 km^2) (e.g., Figure 4.6B), 4% were classified as other terrestrial (28.3 km^2), and 0.3% were classified as water (2.4 km^2) (e.g., Figure 4.6A). This suggests that the majority of damaged areas remained dominated by mangroves despite the

storm damage.

47.6% of damaged pixels had not reached pre-Sidr NDVI levels by mid-2018 (Figure 4.8) with the majority of recovery occurring between 2013 and 2018 (Figure 4.9). Data were not normally distributed so a Spearman's rank test was used to test for correlation between break magnitude and year of recovery. This resulted in a significant trend ($p < 0.01$) with a value of $Rho = -0.24$, indicating a small negative trend (Figure 4.9). The trend remained similar with and without inclusion of the NYR category, which contained nearly 50% of the data.

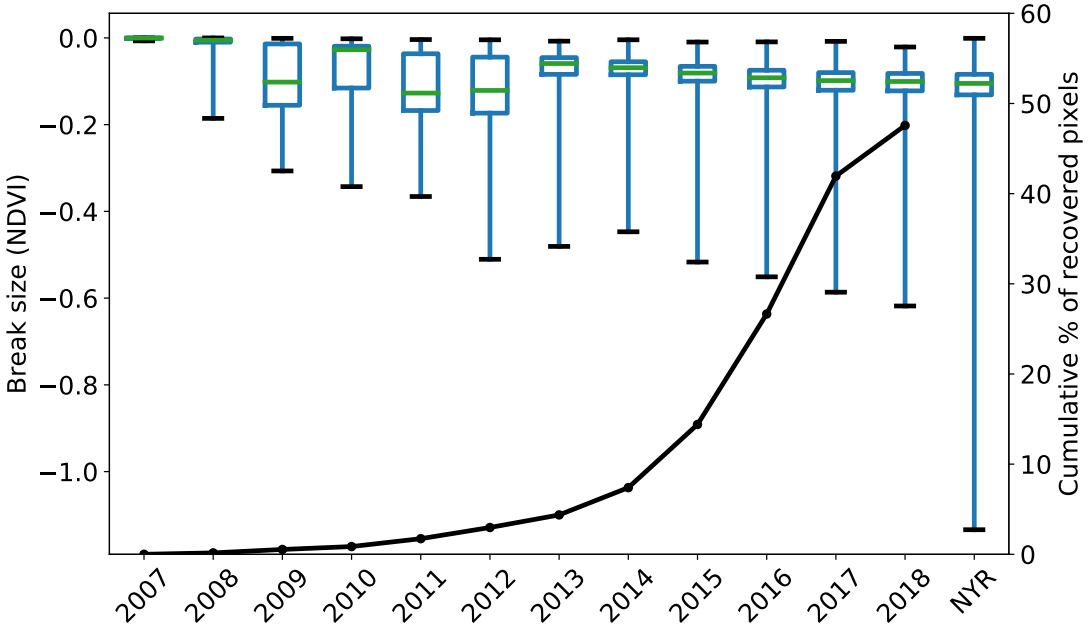


Figure 4.9: Box plot of break size vs. year of recovery (left axis) and cumulative recovery over time (right axis). NYR = Not Yet Recovered.

4.4 Discussion

4.4.1 Classification of Mangroves Using CCDC

The CCDC method proved to be highly accurate for mangrove classification. This demonstrates that CCDC can provide a tractable solution for quantifying change in highly dynamic regions, and can be applied to areas such as the tropics which have

high cloud prevalence. Allocation disagreement was higher than quantity disagreement, indicating more error in the spatial distribution of land cover classes than in the quantity mapped as each class. The main source of error was confusion with other terrestrial vegetation. This is probably because along with other types of tropical vegetation, mangroves do not have a very pronounced seasonal cycle. However, other comparable methods using optical imagery and single-date classifications report overall accuracies in the range of 64–88% [28, 98, 292]. Including temporal information in the classification process in the form of time series models therefore represents a reasonable improvement over single-image approaches. Possible further improvements could be made by including auxiliary data such as elevation or distance from water in the classification step, rather than in post-processing.

The lack of an independent dataset for verification is a limitation of this study. However, CCDC has previously been shown to be a accurate method of determining land cover type [340, 345]. Substantial overlap was also found between the 2010 classification map generated by this study and the GMW estimate, and evidence was found that CCDC was also able to distinguish some smaller areas of mangroves missed by the GMW. Our classification estimates the total area of mangroves in the Sundarbans in the present day to be around 6600 km^2. While the total area of the Sundarbans is often cited as around 10,000 km^2 [91, 97, 130], recent studies using Landsat data place the estimate of actual mangrove forest closer to 6000–7000 km^2 [91, 97, 98]. Our estimate is, therefore, in line with previous studies, and demonstrates that evaluation of baseline extents of important habitats can be more accurate with new techniques, particularly if the full data archive is exploited.

The lower estimates of mangrove extent for 1988 and 1989 indicate that not enough data were available to determine land cover for many pixels. There are two factors which could have contributed to this. Firstly, 1988 was the first year of data used and classifications for this year might have been affected by a spin-up effect of the

model fitting process. If a stable history period could not be initialized by mid-1988 then a pixel would not have a majority class for that year. Secondly, Bangladesh experienced extreme flooding in both 1987 and 1988 [322] exacerbated by a category 3 tropical cyclone making landfall in November 1988 which directly impacted the Sundarbans [155]. This may be why mangrove extent was also lower for 1989. These extreme weather events would have caused rapid changes in land cover and made it more difficult to fit a stable model. The drop in mangrove extent between 2006 and 2007 is discussed in Section 4.4.3.

The classification also indicates that there has been a slight decrease in mangrove extent over the last 20 years. This result is broadly in line with a study by Giri *et al.* [98] which found a slight decrease in mangrove forest in the Sundarbans between the 1990s and the 2000s and a slight decrease in mangrove forest over all of South Asia between 2000 and 2012. Other studies also report little or no long-term decrease in mangrove extent [91, 97]. In both this study and that by Giri *et al.* the detected decrease is within the margin of error of the classification and therefore must be interpreted cautiously. Mangrove dynamics in the Sundarbans are also complex, with both land erosion and accretion contributing to mangrove numbers [11, 98]. While evidence was found that mangrove quantity was increasing before Sidr but decreasing afterwards, the majority of that reduction occurred in the last five years. In addition, the number of pixels classed as non-mangrove after Sidr which were previously classed as mangrove was 4.3%, much higher than the estimated reduction in mangrove extent between 2008 and 2017. This suggests that even where mangroves were damaged badly enough to be classified as non-mangrove for a time, most either eventually re-established as mangroves or the loss was offset by gains elsewhere. This demonstrates the resilience of these habitats to events such as cyclones. However, long-term degradation and subsequent fragmentation could reduce this resilience to critical levels.

4.4.2 Long Term Vegetation Trends in the Sundarbans

Vegetation dynamics in the Sundarbans are complex, particularly when examined over a long period of time. By examining overall greening and browning trends, it was hoped that areas of the Sundarbans at potential risk of degradation in the future would be highlighted. Results showed that three-quarters of pixels identified as mangroves at some point during the 30-year period from 1988 to 2017 exhibited a positive trend over their entire time series. Of these areas, some are clearly regions of land accretion where new stands of mangroves have become established (Figure 4.7). Given that mangrove extent changed little over the study period, many of these areas of accretion must be being counterbalanced by mangrove loss. However, approximately three times more strong positive trends were found than strong negative trends. This suggests that over the majority of the Sundarbans, mangrove canopy cover has increased over the past 30 years. Excluding establishing mangrove stands, mangroves exhibiting a net increase in greenness could fall into one of several categories. Firstly, they could have experienced damage (e.g., from a cyclone) but recovered past their pre-damage greenness level. Secondly, they could have experienced damage prior to 1988. In this case, even if the mangroves did not recover to their pre-damage level, the trend would still present as being net positive. Finally, an increase in NDVI could be the result of a change in mangrove species composition. *C. decandra*, *E. agollocha*, and *H. fomes* often coexist in the Sundarbans, with the shorter *C. decandra* often forming as an understory to *E. agollocha* [91]. Decline in one species could therefore be masked in terms of NDVI trend if another species rapidly takes over and establishes.

Approximately one quarter of pixels classified as mangroves at some point in the 30-year period were found to have an overall negative trend. This suggests that even though there is very little net loss of mangroves, a substantial proportion may be at risk. In particular, there is a clear area to the north where negative trends

are prevalent. A study into long-term mangrove species dynamics by Ghosh *et al.* found that this area has predominately been occupied by *H. fomes*, which has a low tolerance for saltwater. Die-back disease has been a large-scale problem since 1980 and was estimated to affect 5–6% of *H. fomes* in 2010. It is, therefore, possible that a large proportion of the negative trends found were in *H. fomes*. This highlights the particular vulnerability of these important habitats to global climate change and subsequent sea level rise, as well as increasing number of extreme weather events. As such, monitoring mangroves in this way can provide a barometer for regional and global vegetation changes.

4.4.3 Investigation of Dynamics around Cyclone Sidr

4.4.3.1 Comparison to Previous Damage Estimates

Results showed that around 11% of the Sundarbans mangroves were damaged by cyclone Sidr. This is lower than other estimates. However, it is important to highlight that the approach used in this study is substantially different to other studies, which used relatively few images (e.g., [6, 28, 55]). By fitting temporal models rather than relying on single-image classifications [6] or differencing between images [28], mangrove dynamics both before and after the cyclone could be more accurately captured. This creates difficulty in comparing the damage estimates between CCDC and the other methods. In addition, this study covered the entire Sundarbans region, including many outlying areas not covered by other studies. This likely reduced the damage estimate as a percentage.

Akhter *et al.* [6] used Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) imagery from November 2007 to asses damage immediately following the cyclone. Based on field visits, the study area was classified into regions which were highly, moderately, and severely affected by Cyclone Sidr with the total damaged area estimated to be 22.2%. However, this study only focused on a small

area of the forest which was directly impacted. Work by Bhowmik and Cabral [28] classified four Landsat 7 scenes from February 2007–2010 to distinguish between different mangrove species before using NDVI values to track disturbance and recovery from Sidr. This placed the estimate of damage at around 45% of the Sundarbans, with recovery being complete by 2010. However, this assessment was based on more generalized reductions and increases in relative NDVI values across images, and still only covered the area directly impacted by Sidr. Cornforth *et al.* [55] used backscatter values from radar imagery as a proxy for above ground biomass, where a reduction in backscatter indicates a reduction in biomass and therefore a change in forest condition. Imagery from 2007 was compared to images from 2008 and 2009, though due to a lack of available imagery, less than half of the total area of the Sundarbans was assessed. This study reported lower damage estimates than other studies which used optical data. However, their approach was still based on differencing between single images from different years and using a threshold to identify damaged areas. This meant that determination of changed areas was entirely dependent on relative change from a single previous time point. By fitting a stable temporal model to optical data, some of the problems caused by noisy or missing data can be eliminated by capturing the broader seasonal dynamics of a land cover type. Change analysis carried out this way is far less dependent on the value of individual observations.

This study also took a conservative approach to damage estimation by only looking for changes in the two months following Sidr. This was based on the fact that it is difficult to validate what truly caused a change. In addition, there were limitations on allocating the true date of change due to data availability. The only data available between September 2007 to February 2008 for the area directly impacted by Sidr was from the Landsat 7 satellite, and no data at all was available for June, July, or August 2007. This meant that the SLC failure on board Landsat 7 contributed to the lower estimate given by CCDC. While season-trend modelling approaches such as CCDC

can interpolate between observations to fill in gaps in the data record, change cannot be assigned until an actual observation is available. For pixels with missing data due to the SLC failure, no observation was available until March 2008, meaning that if those pixels were damaged by Sidr, the change would not be detected by CCDC until nearly four months after Sidr made landfall. As two months was used as the cut-off for changes to be reliably assigned to Sidr, these damaged pixels were missed leading to gaps in the estimate of damage extent (Figure A.1). This problem was difficult to avoid without substantially extending the range within which a change was considered attributable to Sidr, and therefore decreasing the likelihood of Sidr being the true cause of change. 11% is, therefore, a robust minimum estimate of the area of mangroves damaged by Sidr. Estimates based around comparing yearly images (e.g., [28]) are more likely to attribute changes to Sidr which may have other causes, and to be more affected by noise and cloud contamination.

4.4.3.2 Impact of Sidr on Mangrove Extent

The impact of Sidr contributed to a drop of 90.1 km^2 in mangrove extent between 2006 and 2007 (Table A.1), despite the fact that the vast majority of damaged pixels were classified as mangrove before and after Sidr. This is because classification is based on a majority class in a given year. For some pixels damaged by Sidr, valid data was only available for May 2007 and March 2008. In these cases, a change would be detected in March 2008 once Landsat 5 data was available, with the end date of the last stable model being May 2007. Since this model covers less than half the year, these pixels would not have a majority class for 2007.

The attributed class of most pixels damaged by Sidr did not change, but the damage done was substantial enough to be flagged by the CCDC method. This suggests that the damage done by Sidr was extreme enough to emulate a change in land cover class, but that pixel trajectories after Sidr were still similar enough to other mangrove pixels to still be classified as mangrove. Therefore the CCDC

method is probably capable of further separating the mangrove class into *stable*, *recovering*, and *degrading* mangroves. It is possible that as well as experiencing an increasing or decreasing trend, non-stable mangroves also have more subtle differences in seasonality such as smaller seasonal peaks which would make them distinguishable from the other classes. However, this type of classification would require accurate training data on recovering and declining mangrove populations and was beyond the scope of this study.

Spatially the areas that experienced the highest magnitude of damage were along the coast (Figure 4.8). In particular, the edges of more fragmented areas and smaller islands were more likely to experience greater damage. These areas would have been vulnerable to the storm surge and more likely to be already experiencing erosion or to consist of land which was still in the process of establishing. Given that the majority of breaks were fairly small in magnitude, fairly rapid recovery would be expected. Bhowmik and Cabral [28] investigated the impact of Sidr on biodiversity in the Sundarbans and concluded that the Sundarbans had recovered to a satisfactory level by 2010. Mangroves are also known to recover quickly once primary regeneration has taken place [28] and to be very resilient to many ecological changes [160].

4.4.3.3 Estimation of Recovery from Sidr

Median days disturbed was 80 indicating that the CCDC algorithm was able to fit a new stable model to many pixels within three months. The majority were stable within six months. Given that CCDC is capable of identifying a recovering mangrove pixel just as well as a stable mangrove pixel, these periods of disturbance indicate that many mangroves were damaged enough by Sidr that several months passed before they were once again recognizable as mangroves. This also suggests that fairly small magnitude changes can represent substantial disturbances on the ground.

However, even once new models were fitted, nearly 50% of damaged pixels had not recovered to their pre-Sidr overall NDVI value by mid-2018. While mangrove extent

did drop between 2008 and 2017, the decrease was small, indicating that the vast majority of these pixels remained mangrove throughout that period. The number of pixels recovered increased much more rapidly between 2013 and 2018 than between 2007 and 2013 (Figure 4.9). A possible reason for this could be a protective effect whereby once some mangroves have recovered others are less exposed and recover more quickly.

Another possible explanation is bias in the Landsat 8 NDVI trajectories. There is evidence that Landsat 8 NDVI is positively biased compared to Landsat 5 and 7 data [257, 345]. Zhu *et al.* [345] concluded that this was due to the atmospheric correction method used for Landsat 8. Given that all data was pre-processed using the same method (ARCSI) this was not expected to be an issue. However, the increase in recovered pixels does present around the time of Landsat 8's launch and differences between the sensors cannot be discounted as a possibility. Even if bias exists in the NDVI trends, it is unexpected that so few mangroves appear to have recovered fully by mid-2018. This suggests that although mangroves generally recover quickly, and may appear to have recovered fully based on yearly observations, full recovery following a disturbance event is much slower. If 11% is assumed to be a reasonable minimum area of mangroves damaged by Sidr, the recovery trends indicate that around 5–6% of the Sundarbans mangroves are still recovering from the effects of cyclone Sidr in some way. These mangroves may be more vulnerable to further cyclones and to other environmental changes such as rising sea levels.

4.5 Conclusions

The methods outlined in this chapter provide a tractable solution to exploiting full time series of available Landsat data to map mangrove extent and condition at large scale. Specifically, this was demonstrated for the first time in the Sundarbans, a highly dynamic ecosystem with long term trend changes as well as step changes and

which is in a region prone to persistent cloud cover and associated missing data issues. In this example, mangroves were shown to be resilient in their recovery from large cyclonic events, but the approach was also able to detect longer-term decreases in greenness potentially related to climate change effects such as increased salinity caused by rising sea levels. Monitoring mangroves in this manner can therefore present important evidence as a barometer for climate change. By applying this approach to the the entire Sundarbans region, this study demonstrates that CCDC can provide data to help drive regional or even countrywide scale policies as part of strategies such as REDD+ [200] and blue carbon initiatives, informing future decisions as to where conservation efforts in the Sundarbans and other mangrove ecosystems should be focused.

Chapter 5

Evaluation of the Continuous Monitoring of Land Disturbance algorithm for large-scale mangrove classification

5.1 Introduction

Mangrove forests exist in tropical and sub-tropical regions across the globe, covering a total area of nearly 140,000 km^2 [39, 94]. Growing within the intertidal zone, mangroves occupy a narrow ecological niche and are well adapted to saline environments and harsh coastal conditions [94, 160]. Mangrove forests are of high biological, economic, and ecological importance globally [94, 160], being a source of food, timber, and traditional medicines in addition to aiding in shoreline stabilisation by trapping sediment and nutrients and protecting coasts from the effects of cyclones and tsunamis [23, 172, 188]. Mangrove forests are also important carbon sinks [158], containing on average 1023 Mg per hectare and representing as much as 10% of carbon emissions from deforestation while only accounting for 0.7% of tropical forest globally [68].

Despite their importance, mangrove ecosystems across the globe are under threat [74, 94, 160, 178, 306]. While mangroves are tolerant of environmental change brought by tidal inundation, seasonal rainfall, and storms [178, 212], their ecosystems are vulnerable to the effects of climate change and human activity [74]. As sea levels

rise, inland areas become more frequently inundated with seawater, causing forests to retreat from the shoreline as their tolerance is exceeded [178, 306]. On local scales, conversion to aquaculture (e.g., shrimp farming) is a major threat [160, 236], accounting for around 38% of mangrove forest loss worldwide [236]. In addition, over-exploitation for fuel and timber has resulted in the degradation of a quarter of the world's mangrove forest [236] and mangroves continue to be threatened by increasing human populations, industrialisation, urbanisation, and exploitation of natural resources such as oil [160, 201, 236, 253]. A total of 16% of mangrove species are estimated to be at increased risk of extinction [236]. These pressures, combined with an increasing likelihood of extreme weather events such as cyclones and tsunamis [126], make the assessment and monitoring of mangrove forests a major concern.

This requirement for monitoring is difficult to meet even on regional scales due to the size and diversity of mangrove forests. Mangrove ecosystems are often highly dynamic, with constant erosion and deposition of sediment causing mangrove expansion and retreat, sometimes on very rapid timescales [89]. To accurately monitor these regions requires data collection from the whole forest to obtain a simultaneous snapshot of the entire area. However, for many mangrove forests, field studies are time consuming and expensive due to the size of the area covered and difficulty of access [91, 97]. Remote sensing offers a solution to this problem, with satellites such as the Landsat and Sentinel missions capturing medium-scale spatial resolution data over large tracts of land on weekly timescales. These data have been widely utilised over the last few decades for forest monitoring [33, 53, 200] and for mangrove monitoring specifically [95, 117, 305]. However, while the need for global mangrove monitoring and assessment has long been identified [117, 305], global assessments of mangrove extent and condition remain sparse and are often spatially inconsistent [95].

Several attempts have been made to rectify this knowledge gap. Giri *et al.* used Landsat data to estimate global mangrove extent for the year 2000 [94], but the study

lacked an effective global validation method and did not report classification accuracy. Thomas *et al.* built on this work to compare mangrove extent between 1996 and 2010 using Japanese Earth Resources Satellite (JERS-1) and Advanced Land Observing Satellite Phased Array-type L-band Synthetic Aperture Radar (ALOS-PALSAR) data [286]. Validation was carried out using Landsat data with the study achieving high accuracy. Based mainly on Landsat and ALOS-PALSAR data, the Global Mangrove Watch (GMW) is a collaborative effort and represents the most recent and accurate map of global mangrove extent [39]. In addition to a baseline map for 2010, the GMW has also produced estimates of mangrove change from the mid-1990's onwards. The issues with these efforts are that they either only provide an extent for a single year [94] or provide year-to-year comparisons based on composite imagery, often with large gaps in the change record [39, 286]. While overall accuracy for the GMW is high (94%), the GMW methodology is time consuming, relying on multiple data sets and multiple classification steps in addition to manual quality assurance [39].

Recent developments in the field of land cover monitoring have lead to movement away from year-to-year comparisons of imagery and towards methods which make use of all available observations [338]. The rise of dense time series approaches such as Breaks for Additive and Seasonal Trend (BFAST) [296], Exponentially Weighted Moving Average Change Detection (EWMACD) [35], and Continuous Change Detection and Classification (CCDC) [340] provides a potential solution to the problem of global, long-term monitoring of mangrove extent. These methods use a season-trend modelling approach which captures the seasonal dynamics of land cover while also accounting for changes in condition. Applied on a per-pixel basis, such methodologies reduce reliance on individual observations and on individual cloud-free images, instead relying on the data record for each individual pixel through time. This is especially an advantage for mangrove monitoring, given that mangroves typically grow in tropical and sub-tropical regions with high cloud cover [160].

An additional advantage of these methods is that they allow for the extrapolation of existing data sets. For example, CCDC works by classifying each season-trend model rather than each point in time. A model might cover any time span from a year to multiple decades. By taking an existing, accurate data set such as the GMW, a classifier can be trained based on the models covering the period of that data set. This classifier can then be used to classify all models across time and space, allowing for the generation of yearly and sub-yearly extent estimates without the need for repeating the original, time consuming methodology. This process therefore allows for highly accurate but time-limited data sets such as the GMW to easily be extrapolated through time. In a previous study, it was demonstrated that the CCDC method trained using GMW data could produce highly accurate maps of mangrove extent over the Sundarbans mangrove forest, in addition to tracking changes in mangrove condition over 30 years [17].

In this study, the Continuous Monitoring of Land Disturbance algorithm (COLD) is applied to each of five study sites to produce output detailing when class changes (i.e., mangrove to water) occurred. The COLD algorithm [344] is an update on the previous CCDC algorithm [340] for monitoring class changes in multi-band Landsat time series. The advantage of COLD for tropical forest monitoring is that it operates on a per-pixel basis, taking into account every available observation in the time series and fitting a model to the underlying land cover signal. This removes the requirement for finding cloud-free images for comparison and, since observations from different years are unlikely to fall on the same Day of Year (DOY), the effect of missing data is mitigated when applied to long time series. CCDC has previously been demonstrated to be an effective method for long-term mangrove classification and monitoring [17].

The aim of this study is to demonstrate the applicability of COLD for global mangrove monitoring using the Landsat data archive. Such an approach has potential to produce high quality estimates of mangrove extent which take into account intra-

year fluctuations in addition to providing information on long term loss, gain, and trends. Five study sites are selected from across the globe to represent a range of mangrove species and forest types, with a variety of land cover change drivers. These sites also vary widely in the quantity of available data. Yearly mangrove extent for each site is calculated and tracked over time, with comparison to the GMW and other mangrove extent studies. Validation is performed through expert manual interpretation of Landsat and Google Earth imagery. The efficiency and feasibility of the method applied in Chapter 5 for global mangrove monitoring is demonstrated and discussed.

5.2 Materials and Methods

5.2.1 Study Areas

Five study sites were selected to represent a diverse range of conditions, mangrove populations, and data densities (Figure 5.1). An overview of each site is given here and summarised in Table 5.1.

Table 5.1: Description of the five study sites.

Study site	Path/Row	Rainfall (mm/year)	Species	No. of Scenes
Niger Delta	189/57	3000–4500 [201, 214]	*Rhizophora racemosa*, *R. mangle*, *R. harrisonii* [136, 327]	163
French Guiana	227/57	2000–3000 [10, 79]	*Avicennia germinans*, *Laguncularia racemosa*, *Rhizophora* sp. [79, 240]	203
Borneo	117/58	1800–3000 [48, 315]	*Avicennia* sp., *Sonneratia* sp. [242, 315]	392
Malaysia	128/57	2000–2800 [216, 217]	*Rhizophora apiculata*, *Rhizophora mucronata* [2, 100]	605
Australia	99/72	600–1800 [69]	*Avicennia marina*, *Rhizophora stylosa* [13, 69]	890

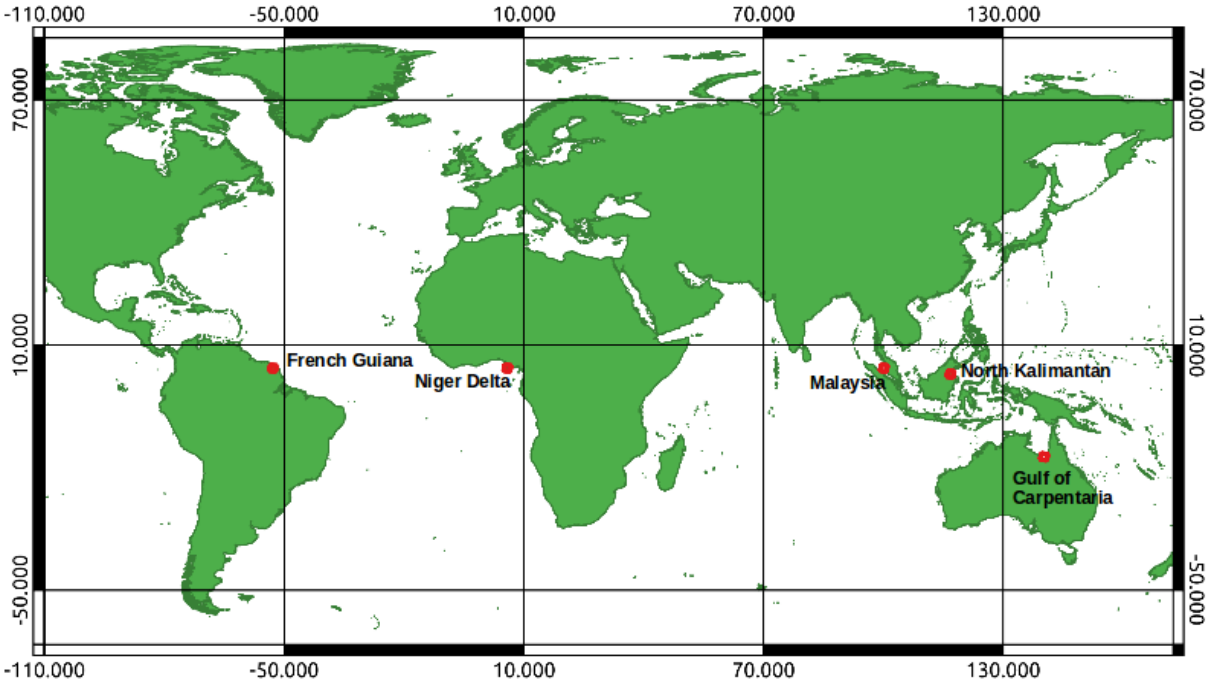


Figure 5.1: Location of the five study sites.

5.2.2 Niger Delta, Nigeria

Draining into the Gulf of Guinea on the coast of southern Nigeria, the Niger River delta is one of the world's largest wetlands covering an area of approximately 70,000 km^2 [201]. It contains Africa's largest contiguous mangrove forest [212] and the third largest mangrove forest globally [159]. The dominant mangrove species in the area are *Rhizophora racemosa*, *R. mangle*, and *R. harrisonii* with *Avicennia germinans* and *Laguncularia racemosa* also present [136, 327]. The Niger Delta is also Africa's largest river delta and is home to around 20% of Nigeria's population [159]. The area used in this study is located from 3°59′ to 5°4′N and 5°14′ to 6°36′E (Figure 5.2A), covered by Worldwide Reference System-2 (WRS-2) Path 189 Row 57 (Table 5.1).

The region experiences very high annual rainfall (Table 5.1), with two peaks in July and September, a short dry season in August, and a longer dry season from October to March [201]. The delta is rich in biological resources, with mangroves being used for fuel, wood, fish trapping, local craft and construction. However, there is

concern that local populations are over-dependent on the mangrove forest, especially for fuel wood [201]. The Niger Delta also has large oil and gas deposits which have been heavily exploited for decades [159], contributing substantially to mangrove forest loss [201]. The expansion of the oil industry typically involves the creation of canals for exploration and access, causing contamination of freshwater systems with seawater and destroying local ecosystems [132]. In addition, since the 1970's there have been multiple unrecovered oil spills [132], causing widespread damage to the mangrove forest [159]. Estimates suggest that between 9 and 13 million barrels of oil have been spilled in the Niger Delta since 1958 [22].

Despite the ecological significance of the Niger Delta, mangroves and the potential negative effects of oil and timber exploitation, the region is generally under-studied. While several studies have been carried out which utilise remote sensing [132, 159, 212, 214], these have been limited to comparisons between single images, offering change analysis on a five to ten year scale. This limitation can be attributed to the high cloud cover in the region, which makes it difficult to find suitable images for comparison (Figure 5.3).

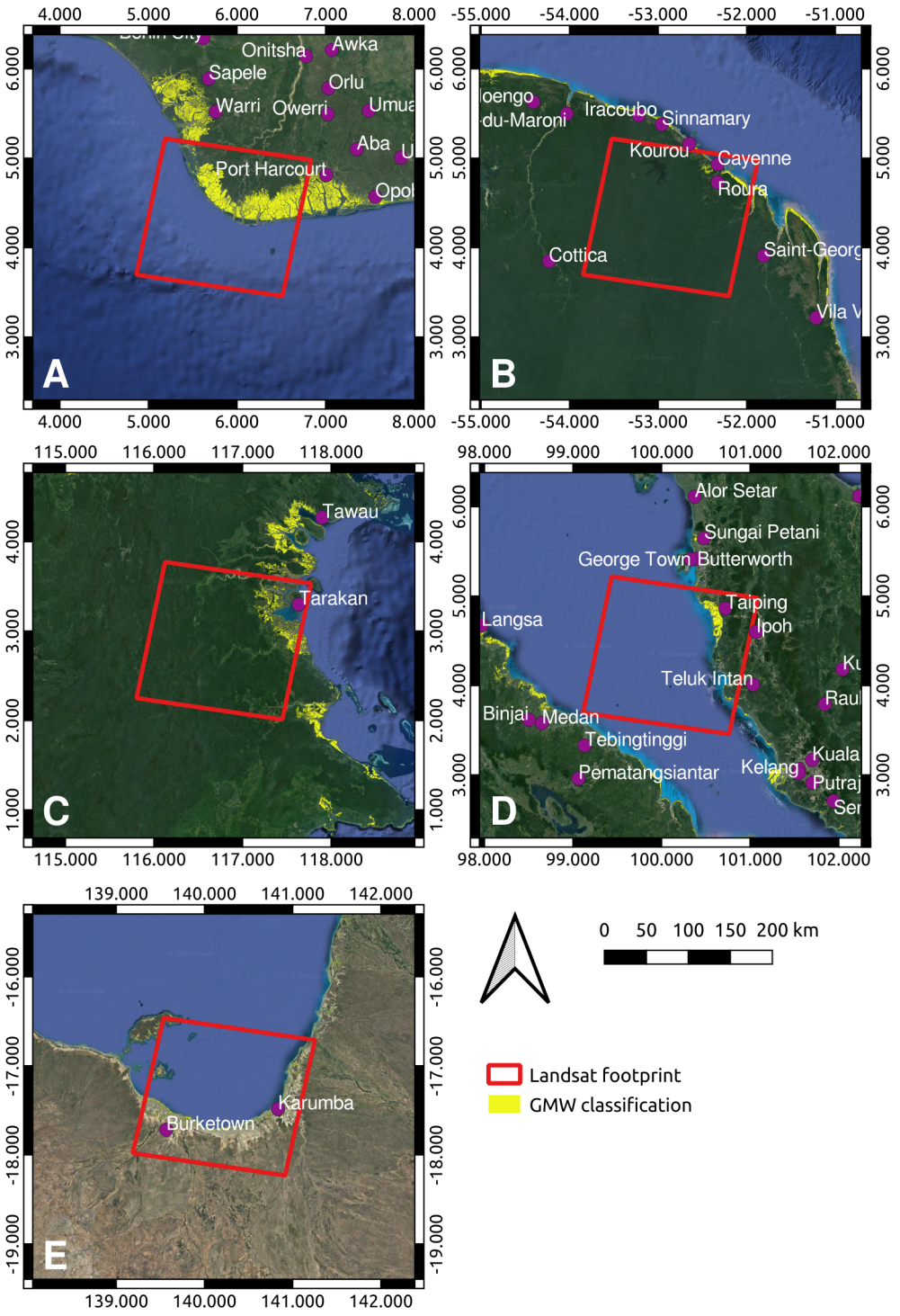


Figure 5.2: Overview of each of the five study sites. (**A**) Niger Delta, Nigeria (**B**) French Guiana (**C**) North Kalimantan, Borneo island (**D**) Matang Forest Reserve, Malaysia (**E**) Gulf of Carpentaria, north Australia. GMW = Global Mangrove Watch.

5.2.3 Area around Cayenne, French Guiana

At 1500 km, the coast of north eastern South America is the longest muddy coast in the world [10]. Stretching from the mouth of the Orinoco to the mouth of the Amazon river, the coastal geology in this region is dominated by the large-scale deposition of sediment from the Amazon [10, 89]. This particulate discharge forms mud banks up to 60 km long and 30 km wide, which migrate along the coast of French Guiana towards the Orinoco at a rate of up to 3 km per year [87]. This migration creates a particularly unstable and dynamic coastline, with rapid and constant erosion and accretion [10, 89]. The climate is hot and humid year round, with high levels of rainfall.

Mangroves cover an area of approximately 700 km^2 across the coastline of French Guiana [79]. However, there is a complex relationship between mud bank migration and mangrove advance and retreat along the shoreline which makes mangrove dynamics in this region difficult to predict in the long term [87]. Typically, the presence of mud banks leads to the formation of large intertidal mud flats which are colonized by mangroves before the onward migration of the mud banks causes a loss of protection followed by massive erosion and mangrove retreat [87, 89]. As a result, mangroves along the French Guianese coastline can be characterized into four development stages which exist roughly parallel to the coastline, from early growth and establishment to declining and dead mangrove forest [240]. *Avicennia germinans* is the dominant species, forming pioneer stands alongside *Laguncularia racemosa* in addition to making up the majority of the adult-age mangrove stands, with stands increasing in age with distance from the sea [79, 240]. Other species present include *Laguncularia racemosa* and *Rhizophora mangle*, the former being a pioneer species which grows close to the shoreline, whereas the latter is less salt-tolerant and grows further inland [79].

The area used in this study is located from 3°24′ to 5°16′N and 51°57′ to 53°42′W (Figure 5.2B), covered by WRS-2 Path 227 Row 57 (Table 5.1). While studies have been carried out in this region which utilise satellite data [79, 87, 301], there are no current studies which make full use of the Landsat archive to monitor mangrove extent over time.

5.2.4 North Kalimantan, Borneo Island

An area rich in biodiversity, Borneo is the third largest island in the world [88]. Kalimantan is the Indonesian part of the island and represents approximately three-quarters of its area, with the remaining quarter being split between Malaysia and Brunai [161]. The climate is characterised by frequent rainfall and high temperatures year round [161]. In addition to mangroves, Borneo's forests include dipterocarp, freshwater, peat, swamp, and heath forests, as well as *nipah* or mangrove palms, which grow in coastal regions alongside mangroves [88]. Specific information on mangrove species composition in North Kalimantan is difficult to find, but the area is likely dominated by *Avicennia* sp. amd *Sonneratia* sp. [242] with *Rhizophora apiculata* also being present [315].

Estimates suggest that in the early part of the 20th century Borneo was dominated by forests, which covered around 75% of the island; however, about half of this forest has since been lost to deforestation [48]. Gaveau *et al.* [88] found that between 1973 and 2010, deforestation in Borneo occurred at twice the rate of any other humid tropical forest, primarily due to the expansion of industrial scale oil palm plantations. This rapid deforestation has been found to be a significant contributor to rising temperatures in the region due to the loss of the evaporative cooling effect of the forest canopy [48]. In Kalimantan, the biggest threat to mangrove forest is from conversion to aquaculture for fish and shrimp farming [253].

The area used in this study is located from 2°0′ to 3°47′N and 115°55′ to 117°39′E

(Figure 5.2C), covered by WRS-2 Path 117 Row 58 (Table 5.1). While previous studies have not focused on this area specifically, there is evidence of major deforestation in this part of North Kalimantan [88, 253] which could be further investigated using COLD.

5.2.5 Matang Forest Reserve, Malaysia

The Matang Forest Reserve (MFR) is located on the north west coast of peninsular Malaysia in the state of Perak. The region has a warm and humid equatorial climate with two monsoon seasons, one between November and March, and one between May and September [217]. Covering approximately 40,000 ha [100, 127], the MFR has been a managed forest for over 100 years producing charcoal and timber [217]. *Rhizophora apiculata* and *Rhizophora mucronata* make up around 80% of mangroves in the area, being the main commercially grown species [2]. Other species present include *Avicennia officinalis*, *Sonneratia alba*, *Bruguiera cylindrica*, *Bruguiera parviflora*, *Ceriops tagal*, and *Excoecaria agallocha* [2, 100, 127].

Timber extraction occurs in about 80% of the mangrove forest [1] and is carried out in a 30-year cycle [100]. Blocks of trees are thinned after 15 and 20 years of growth to provide wood and allow more space for the remaining trees [100]. After 30 years the block is clear felled for charcoal and replanted [100]. This approach is widely considered to be a sustainable form of silviculture [1, 100, 127].

The area used in this study is located from 3°23′ to 5°10′N and 99°48′ to 101°3′E (Figure 5.2D), covered by WRS-2 Path 128 Row 57 (Table 5.1). This area covers the entirety of the MFR. Because of the quantity of both current and historical inventory data available, the MFR has been extensively studied using remote sensing methods. Due to the timescale involved, the majority of studies have utilised Landsat imagery [1, 2, 127, 213, 216, 217] though data from Unmanned Aerial Vehicles (UAVs) have also been used [218].

5.2.6 Gulf of Carpentaria, Australia

The Gulf of Carpentaria (GOC) is located on the north coast of Australia in Northern Queensland and the Northern Territory. The climate in this region is hot and humid with rainfall concentrated in the wet season, which lasts from December to March. Rainfall during this period often exceeds the capacity of river systems causing widespread flooding [13]. A narrow strip of mangroves exists along much of the Gulf's coastline and has remained relatively undisturbed for the last two centuries [13]. Change in the region is therefore likely to be caused by natural events or by the indirect effects of anthropogenic climate change rather than by direct human activity [13]. These mangroves play a vital role in the local ecosystem, providing nurseries for a variety of aquatic life and protecting coral reefs and sea grass [112] in addition to protecting the shoreline from the impact of cyclones [176], which occur two to three times a year [13]. More than 30 mangrove species are reported to grow in the region [13, 69], though *Avicennia marina*, *Rhizophora stylosa*, and *Sonneratia alba* are predominant [13]. Mangroves in the GOC experienced an extreme dying event in 2015, with the loss of over 7000 ha of mangroves [69, 112]. This die-back event has been attributed to cumulative stress due to climate factors including lower than average sea levels and rainfall [69, 112].

The area used in this study is located from 16°27′ to 18°19′S and 139°16′ to 141°10′E (Figure 5.2E), covered by WRS-2 Path 99 Row 72 (Table 5.1). This footprint covers the southernmost region of the GOC including the mouths of the Albert, Leichhardt, Flinders, Bynoe, and Norman rivers, all of which frequently discharge large quantities of water and sediment into the Gulf [13]. These large seasonal changes in combination with the ongoing effects of climate change on the GOC make the area well suited to long-term studies using Landsat data [13, 182].

5.2.7 Data and Pre-Processing

5.2.7.1 Landsat Data

All United States Geological Survey (USGS) Collection 1 Tier 1 Landsat 4-5 TM, Landsat 7 Enhanced Thematic Mapper Plus (ETM+) and Landsat 8 Operational Land Imager (OLI) data covering each study site for the period from January 1989 to June 2020 were downloaded from the Google public repository. Scenes with greater than 90% cloud cover were not included on the basis that cloud contamination of the remaining area is highly likely. Collection 1 Tier 1 provides high quality data which have been georegistered and inter-calibrated across the Landsat instruments, and are considered suitable for time series analysis [291]. The downloaded images were atmospherically and radiometrically corrected and converted to analysis ready surface reflectance data using the Atmospheric and Radiometric Correction of Satellite Imagery (ARCSI) Python package [38]. As part of this process, cloud masks were also created using the Function of mask (Fmask) algorithm [341].

After processing to an analysis ready format all data were uploaded to the Supercomputing Wales (SCW) platform and indexed into a data cube for ease of analysis. The Open Data Cube (ODC) is an open source initiative which utilises a PostgreSQL database along with a Python interface to simplify the processing, organisation, and access of geospatial data [18]. Data are spatially aligned to allow for per-pixel analysis through time and processing can easily be parallelised to facilitate analysis of large scale data sets [165]. All scenes were left in their native resolutions and coordinate systems. Based on work by Zhu et al. [342] and the results of Chapter 4, only the red, green, near-infrared (NIR), and shortwave infrared (SWIR) bands were selected for analysis.

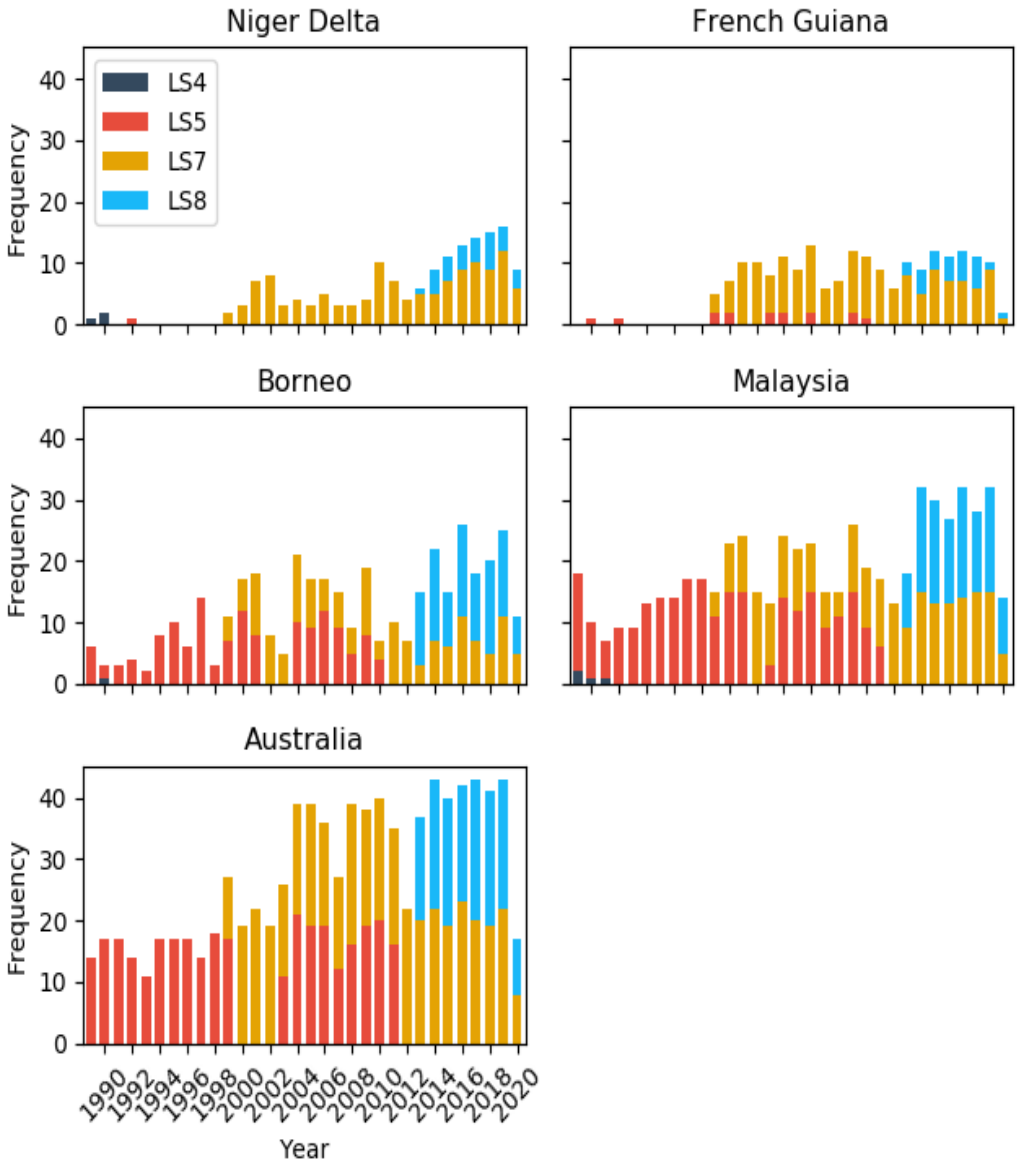


Figure 5.3: Plots showing the number of scenes downloaded from each satellite for each of the 30 years. There is a clear lack of imagery for the 1990's for the Niger Delta and French Guiana sites, and a heavy reliance on Landsat 7 imagery. Data quantity for 2020 is lower because data were only processed up to mid-2020. LS4 = Landsat 4, LS5 = Landsat 5, LS7 = Landsat 7, and LS8 = Landsat 8.

There was a large disparity between the number of scenes available for each site, with the Niger Delta having the fewest available images and the Gulf of Carpentaria having the most (Figure 5.3). No data were available for the Niger Delta or French Guiana sites for 1991 and 1993–1998. Both of these sites were also highly dependent

on Landsat 7 imagery, with only one Landsat 5 image being available for the Niger Delta over the entire time series and fewer than 20 being available for French Guiana (Figure 5.3). Of the scenes that were available for the Niger Delta, French Guiana, and North Kalimantan, most were highly contaminated with cloud cover.

5.2.7.2 Auxiliary Data

Global Mangrove Watch Baseline As in Chapter 4 the Global Mangrove Watch (GMW) 2010 mangrove map (Version 2.0) was used as a basis for the classification. This map provides a highly accurate global mangrove baseline for 2010 [39].

Elevation The Shuttle Radar Topography Mission (SRTM) global 1 arc second product was downloaded and elevation data were extracted for each study site.

Land/Water Masks Land/water masks were generated for 2010 to separate non-mangrove areas into land and water classes. This was done by calculating the mean Normalized Difference Water Index (NDWI) for each pixel between January 2009 and December 2011. Three years of data were used because for the Niger Delta and French Guiana sites, insufficient data were available to generate a mask value for every pixel using 2010 alone. NDWI utilises the green and NIR bands for water detection in wetland environments [194]. Typically, NDWI values above zero represent water while values below zero represent land [193, 194]. However, given tidal and river effects, a value of -0.3 was found to be more effective across all of the study sites.

Distance to Water Rasters Once land/water masks had been generated for each site, the *gdal_proximity* function from the Geospatial Data Abstraction Library (GDAL) [83] was used to generate distance-to-water rasters for each site.

5.2.8 Classification of Mangroves Using COLD

Classification maps identifying pixels as either mangrove, other terrestrial, or water were generated for each year from 1989 to 2019 (inclusive). The classification process used to generate the yearly class maps is described in detail in Chapter 4. The methodology will be summarised here.

5.2.8.1 The COLD Algorithm

COLD works by fitting a linear seasonal model to a stable history period then comparing new observations to the existing model. The general form is the same as for CCDC and is described in Equation (5.1), where $\hat{\rho}(i, x)$ is the predicted value for the ith Landsat band at Julian date x, $a_{0,i}$ is the coefficient for the mean of the ith Landsat band, $a_{1,i}$ and $b_{1,i}$ are coefficients representing intra-annual change, and $c_{1,i}x$ is the coefficient representing the inter-annual change, or trend [340]. $T = 365.25$ (the number of days in a year). Change is assigned when it is clear that the new observations do not fit the existing model. COLD therefore outputs a set of model parameters, each covering a specific period of time. These parameters can then be used as input for a classifier to classify each time segment [340, 344].

$$\hat{\rho}(i, x) = a_{0,i} + a_{1,i}cos\left(\frac{2\pi}{T}x\right) + b_{1,i}sin\left(\frac{2\pi}{T}x\right) + c_{1,i}x \tag{5.1}$$

COLD updates the original CCDC algorithm with the goal of reducing errors. To achieve this a more robust change detection method is developed by Zhu *et al.* in [344]. Rather than using a threshold of three times the RMSE of the model to identify change, COLD takes advantage of the fact that the sum of the squared model residuals follows a chi-squared (χ^2) distribution, where the number of Degrees of Freedom (DOF) is equal to the number of spectral bands. Using a Percent Point Function (PPF) with a value of 0.99, a threshold for change can be calculated using Equation (5.2), where $\rho(i, x)$ and $\hat{\rho}(i, x)$ are the actual and predicted values for Landsat band

i at time x, and k is the number of Landsat bands used [344]. If this threshold is exceeded six times consecutively, a potential change is flagged [344]. Observations falling within the threshold are added to the existing stable period and included in the current model [344].

$$\sum_{i=1}^{k} \left(\frac{\rho(i,x) - \hat{\rho}(i,x)}{RMSE_i} \right)^2 \chi^2(k) > \chi^2_{0.99} \tag{5.2}$$

In addition, COLD uses the angle between consecutive change vectors to confirm change. The basis of this method is that a true disturbance is likely to be consistent in its direction. For each anomalous observation i, the angle between the change vectors $\beta_{i,i+1}$ is calculated (Equation (5.3)). Zhu *et al.* [344] suggest that a mean angle of less than 45° between consecutive change vectors indicates a persistent change, i.e., that observations after the supposed date of change continue consistently along a new trajectory. To discount changes due to regrowth, an additional step utilises the red, near-infrared (NIR), and shortwave infrared (SWIR) bands: if NIR is increasing but red and SWIR are decreasing, land cover is becoming greener, possibly indicating regrowth rather than a class change [344]. These regrowth breaks can be removed by checking whether the rate of green-up was faster or slower before the break. If it was faster, the break is likely to be due to regrowth which has stabilised [344]. Inclusion of these steps reduces commission error from both regrowth events and ephemeral change [344].

$$\frac{1}{k} \sum_{i=1}^{k-1} \beta_{i,i+1} < 45^\circ \tag{5.3}$$

COLD was implemented in Python 3.6.8 and run over each study site using the SCW facility to speed up processing. COLD uses Lasso regression, which minimises overfitting through regularisation by limiting the magnitude of the model coefficients [342]. The degree of regularisation is controlled by a parameter λ where $0 < \lambda < \infty$.

A value of $\lambda = 0.1$ was chosen for this study as visual assessment suggested that this produced better delineation between mangroves and water than $\lambda = 1$.

However, to further speed up runtime the re-initialisation period of the model was increased from one day to 90 days. This is the length of time to allow, before updating the model, for inclusion of any new stable observations [344]. Updating the model is a computationally expensive process and adds significantly to algorithm runtime. A minimum period of one year is recommended between model updates; however, this can increase commission error because it reduces the adaptive capability of the algorithm [340, 344]. The 90-day re-initialisation period was chosen as a compromise between runtime and change detection accuracy. COLD was successfully run over the five study sites in less than four days using the Supercomputing Wales (SCW) platform, which provided access to up to 600 cores for processing (though not all cores were continually available). This represents an improvement of around 30% over the previous study which took two weeks using the same computational resources [17].

5.2.8.2 COLD Outputs

For each model covering a specific time period, the outputs from COLD were the per-band model coefficients as described in Equation (5.1), RMSE, and an overall value for each model calculated using the slope and intercept [340]. Given an input of five spectral bands and a third-order harmonic model, this resulted in a set of 45 variables for each model.

5.2.8.3 Model Training

The outputs produced by COLD were used to classify each model (and therefore each stable time period) as mangrove, water, or other terrestrial using the method provided by Zhu *et al.* [340] and previously implemented in Chapter 4. As different land cover types have different seasonal cycles, a classifier can be trained to identify different land cover types based on the model coefficients and other COLD outputs.

This requires the generation of a training set of these model outputs, where the land cover class represented by each model is known. Once trained, the classifier can then be used to classify all models across space and time.

To generate a training data set, the Global Mangrove Watch (GMW) 2010 mangrove/non-mangrove mask was combined with the land/water mask for each of the five study sites to create a map with each pixel masked as being mangroves, other terrestrial, or water. A total of 500,000 sample pixels were then randomly selected for each class, for each of the locations, giving 1.5 million samples per site and 7.5 million samples overall. The ability to produce large training data sets incorporating a wide range of intra-class variation is an advantage of this method. The sample size was based on the assumption that a 1% sample is desirable (given that a single Landsat scene consists of approximately 50 million pixels) and that some samples would have to be discounted as having unstable land cover. Stable in this case means that for that location, the land cover class as taken from the combined 2010 mask remained the same for the entirety of 2010, i.e., a stable model existed for that location which began before January 2010 and ended after December 2010. Models covering a period starting before 2010 and ending after 2010 likely represent a land cover signal as defined by the 2010 data, and are therefore suitable for inclusion in the training set. Using this method the COLD model outputs for each location were checked for stability and added to the training data set if the land cover were stable for 2010.

The sizes of the final training data sets are given in Table 5.2.

Table 5.2: Number of training samples used for each site, for each class. A total of 500,000 random samples were taken for each class, before only the stable models for the training year (2010) were screened out and used to train the classifiers.

Study Site	Other Land	Water	Mangroves	Total
Niger Delta	217,348	422,090	413,954	1,053,392
French Guiana	310,546	477,790	204,861	993,197
Borneo	315,859	497,380	269,027	1,082,266
Malaysia	248,994	499,396	465,032	1,213,422
Australia	91,057	457,280	227,171	775,508
Total	1,183,804	2,353,936	1,580,045	5,117,785

For each sample, the 45 COLD outputs for the period covering 2010 were used to train a set of Random Forest classifiers. Five classifiers were trained, one individual classifier per-site and one classifier which was trained over the data from all sites. This was done in order to compare whether site-specific classifiers would be more accurate than a classifier trained over the global data set. In all cases the samples were randomized before training and an 80/20 train/test split was used to assess the training accuracy of the classifier. All classifiers were implemented using the scikit-learn Python library [227].

5.2.8.4 Generation of Yearly Class Maps

Once the classifiers were trained, all models produced as output from running the COLD algorithm per-pixel were classified twice: once using the overall classifier, and once using the site-specific classifier. This resulted in a set of classified models for each pixel, each with a start and end date. While some models covered the entire 30-year time period, others only covered a few years. To summarise these data into yearly class maps, pixels were assigned the majority class within a given year. If no majority existed, the pixel was assigned a value of 0 (not enough data) for that year. Mangrove pixels were given a value of 1, other terrestrial pixels a value of 2, and water pixels a value of 3. Class maps for 1989 and 2020 were not generated. The

year 1989 was excluded because, in a previous study, the classification for the first year in the time series was found to be unreliable due to spin-up effects of the COLD algorithm [17]. The year 2020 was excluded because data were only processed up to mid-2020 and therefore not enough data were available to decide a majority class for that year.

5.2.9 Post-Processing

The generated class maps for 1990-2019 were first cropped to the same footprint to ensure consistency of extent over time. This was necessary because the different Landsat satellites have different footprints, leading to slight differences in the extent of maps from different years. Mangroves only grow at low elevations in tidal and intertidal zones [98], so any pixels classified as mangroves at above 30 m elevation or more than 2 km from a water body were assumed to be miss-classifications and removed using the SRTM products and distance-to-water rasters, respectively. Finally, pixel clumps less than 900 m^2 (0.9 ha) in area were removed and replaced with the largest neighbouring class to reduce error related to small-scale features [39]. Yearly mangrove extent in km^2 was then calculated for each site based on the number of pixels in each map assigned to class 1. Manual quality assurance was then carried out to remove any remaining erroneous pixels.

5.2.10 Validation

Validation using field data or a separate satellite data set was not possible due to the spatiotemporal extent of the study. There is no other dataset which covers the same time period as the Landsat missions at the same or higher spatial resolution. Classifier validation was therefore performed using the same methodology as the GMW [39] and previously used in Chapter 5. For each site, 15 random years were selected out of years with available imagery (e.g., No data were available for the Niger Delta site for much of the 1990's, meaning no validation samples could be taken for those years). For each

year, 200 validation points were selected for each class using stratified sampling, with the class being determined by the generated mangrove/water/other terrestrial map used to generate the training data. Stratification was used to ensure that sufficient samples were taken from along coastlines where mangroves are predominant. This provided a dataset of 3000 samples for each class for each site (9000 samples total per site, 45,000 overall), randomly distributed through space and time within each stratum. For each validation year, a random scene was then selected from that year to be used for validation. If the selected scene was of poor quality due to cloud cover, another scene was randomly selected until enough data were available to validate at least 50% of the area. Imagery was displayed as an RGB composite of the NIR, SWIR1, and Red bands, which highlights mangroves as spectrally distinct from other vegetation (e.g., see Figures 5.4–5.12). Reference was also made to up to date high resolution Google Earth imagery to aid in mangrove identification. Each validation pixel was assigned a class (mangrove, water, or other terrestrial) through manual human interpretation. While this method of validation results in less certainty than ground-based measurements, it allows for a very high quantity of validation pixels to be generated across the entire spatiotemporal extent of the study site. Where there was uncertainty about the classification for a pixel, the highest likelihood class was assigned based on spatial and temporal context. Uncertain pixels represented a small proportion of the entire validation set for each site, and were therefore thought to be unlikely to influence the final classification accuracy.

Classification with COLD resulted in some pixels with no majority class for any given year. In combination with persistent cloud cover, this meant that not all of the selected points could be used for validation and the final number of validation points used was 34,967. Once validation was complete the User's and producer's accuracy were calculated for each of the three classes along with overall accuracy. Quantity disagreement and allocation disagreement were calculated as described by Pontius

and Millones [237].

5.3 Results

5.3.1 Classification of Mangroves Using COLD

Six classifiers were trained: one individual classifier for each of the five sites trained only on the data for that site, and one overall classifier trained over all of the data. Training accuracy for all classifiers was $> 99\%$ and testing accuracy was $> 95\%$.

Based on the validation carried out, the by-site classifiers and the overall classifier gave essentially identical results, with both producing an overall accuracy of 92.7%. Looking specifically at mangroves, user's accuracy for the by-site classifiers across all sites was 76.4% vs. 77.0% for the overall classifier, and producer's accuracy was 93.4% vs. 92.3%. Values for kappa, quantity disagreement, and allocation disagreement were identical (0.86, 0.05, and 0.02, respectively). Given the similarity of performance and the convenience of having a single classifier, the results presented here are for the overall classifier (Table 5.3).

The class maps generated using the overall classifier were used to calculate and track mangrove extent between 1990 and 2019. For some years and study sites a large proportion of the generated class maps had no valid class (value of zero). This was probably due to a combination of high cloud cover meaning little data were available for those periods and land cover disturbance meaning COLD was unable to fit stable models to those pixels. Maps with less than 60% valid pixels were discounted as being too unreliable and those years were not included in analyses of extent over time.

$$\frac{2 * No.\ of\ true\ positives}{2 * No.\ of\ true\ positives + No.\ of\ false\ positives + No.\ of\ false\ negatives} \quad (5.4)$$

Due to cloud cover, not all points could be validated even when additional imagery was used. As a result, the total number of validation pixels used to calculate accuracy varies by site. The number of mangrove pixels used in validation was also lower than

for the other two classes, especially for the French Guiana and Gulf of Carpentaria sites. This is because mangroves tend to exist in much smaller proportions than the other two classes in addition to being clustered in one area, making it more difficult to find adequate validation data, particularly in images with high cloud cover and/or where land cover boundaries are indistinct. For this reason, the Dice score is also included for the mangrove class (Equation (5.4)) [67, 271], to give a more balanced metric. The overall classifier achieved a Dice score of 0.84 for the mangrove class.

Table 5.3: Results of the spatiotemporal accuracy assessment for COLD over the five study sites, for the overall classifier.

		Reference				
		Mangrove	**Water**	**Other**	**Total**	**User's (%)**
Classifier	Mangrove	3398	153	863	4414	77.0
	Water	156	14,727	1448	16,331	90.2
	Other	128	56	13,627	13,811	98.7
	Total	3682	14,936	15,938	34,556	
	Producer's (%)	92.3	98.6	85.5		**92.7**

5.3.1.1 Niger Delta, Nigeria

When applied to the Niger Delta, the classifier achieved an overall accuracy of 98.1% with a 99% confidence of being between 97.7% and 98.6%. Kappa was calculated to be 0.97, indicating strong agreement between predicted and actual classes, with a Dice score of 0.95 for the mangrove class. Quantity disagreement was 0.007 and allocation disagreement was 0.01. There was a small amount of confusion between mangroves and water and between mangroves and other terrestrial vegetation (Table 5.4). On visual inspection this was mainly caused by over-estimation of mangroves around land cover boundaries, where the line between mangroves other land cover types can be very difficult to define.

When compared to the GMW classification for 2010, COLD estimated mangrove area to be 3168.9 km^2, compared to the GMW estimate of 2616.6 km^2. A total of

19.2% of the area classified as mangroves by COLD was not classified as mangroves by the GMW (Figure 5.4). A total of 2.1% of the area classified as mangroves by the GMW was not classified as mangroves by COLD. Given the high overall classification accuracy this suggests that the GMW substantially underestimates mangrove extent for the Niger Delta.

Table 5.4: Results of spatiotemporal accuracy assessment for the Niger Delta.

		Reference				
		Mangrove	**Water**	**Other**	**Total**	**User's (%)**
Classifier	Mangrove	1026	29	41	1096	93.6
	Water	3	2537	2	2542	99.8
	Other	24	12	2131	2167	98.3
	Total	1053	2578	2174	5805	
	Producer's (%)	97.4	98.4	98.0		**98.1**

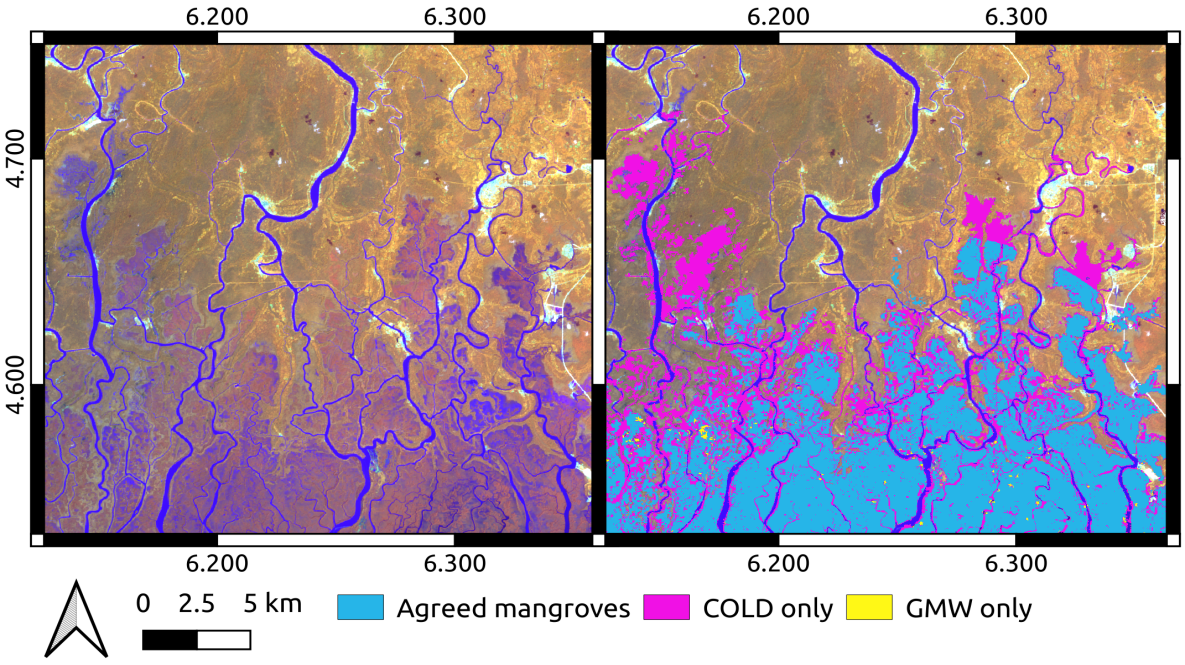


Figure 5.4: Example classification of mangroves in the Niger Delta. Left: false colour Landsat 8 image from December 2013 where Red = NIR, Green = SWIR1, and Blue = Red. Mangroves are clearly visible as having a purple hue which is distinct from the surrounding land cover. Right: classification of mangroves showing areas identified as mangroves by the Continuous Monitoring of Land Disturbance (COLD) algorithm which were missed by the GMW.

Years 1990–1999 inclusive were excluded from extent analysis for having too few valid pixels. For the remaining years (2000–2019), the lowest extent was in 2000 (1806.7 km^2) and the highest extent was in 2009 (3171.6 km^2). Of the 19 years, 9 showed a gain in mangrove extent from the previous year and 10 showed a loss. The Niger Delta experienced the largest loss and gain of any site, with an increase in extent of 704.6 km^2 (39.1%) between 2000 and 2001 and a loss of 699.6 km^2 (25.0%) between 2018 and 2019 (Figure 5.5). While classification accuracy was high for the Niger Delta, the 2000/2001 gain does correspond to some extent with the change in data availability brought by the launch of Landsat 7 (Figure 5.3). Given that extent remained relatively stable between 2003 and 2018, it seems likely that this initial gain is an artefact of data quantity rather than a true change in land cover.

The drop in extent of 25.0% between 2018 and 2019 is partially accounted for by a change in valid data quantity in the class map for 2019. Invalid pixels are those for which no class could be assigned by COLD for that year. While the quantity of invalid pixels in each year remained relatively stable at around 8-13% throughout the 2000s and 2010s, it increased to 22.8% in 2019. A possible explanation could be that, because data availability for the region is so poor, pixels which underwent change in 2019 could not be classified because not enough data remained in the time series to fit a new land cover model. However, the change in valid pixel quantity does not fully explain the drop in extent between 2018 and 2019, and it is likely that the decreasing trend seen since 2013 continued into 2019.

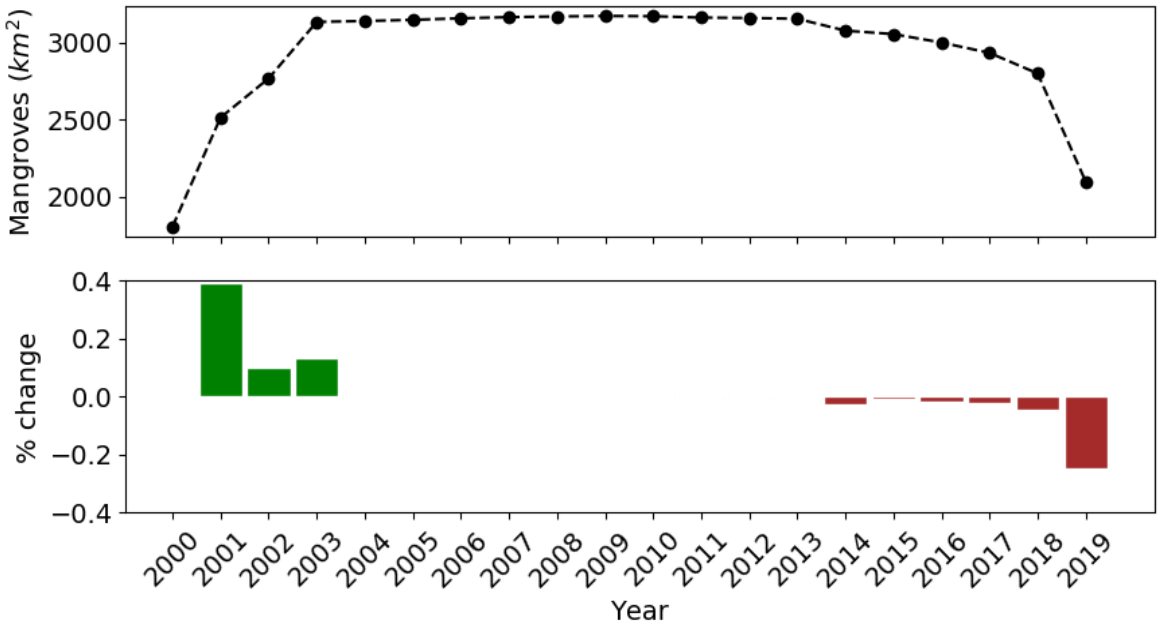


Figure 5.5: Mangrove extent over time for the Niger Delta site. Years 1990–1999 (inclusive) were excluded from analysis because the class maps for those years contained less than 60% valid data.

5.3.1.2 Area around Cayenne, French Guiana

For French Guiana the classifier achieved an overall accuracy of 96.0% with a 99% confidence of being between 95.4% and 96.7%. Kappa was 0.91 with a quantity disagreement of 0.02 and allocation disagreement of 0.02, indicating strong agreement. However, there was substantial confusion between the mangrove and other terrestrial classes with a User's accuracy of 61.2% for the mangrove class (Table 5.5), indicating that around 40% of the pixels identified as mangroves by the classifier were actually in the other terrestrial class, most likely other tropical vegetation. The Dice score for the mangrove class was 0.73. Visually, the boundary between mangroves and other vegetation is very difficult to define in this region as the spectral properties are similar (Figure 5.6).

Table 5.5: Results of spatiotemporal accuracy assessment for French Guiana.

		Reference				
		Mangrove	**Water**	**Other**	**Total**	**User's (%)**
Classifier	Mangrove	504	35	285	824	61.2
	Water	16	3272	20	3308	98.9
	Other	34	8	3330	3372	98.8
	Total	554	3315	3635	7504	
	Producer's (%)	91.0	98.7	91.6		**96.0**

For 2010, COLD estimated mangrove area to be 185.1 km^2, compared to the GMW estimate of 163.6 km^2. 28.7% of the area classified as mangroves by COLD was not classified as mangroves by the GMW, whereas 19.3% of the area classified as mangroves by the GMW was not classified as mangroves by COLD (Figure 5.6). While extent estimated from the two methods is similar, this indicates that there is substantial disagreement in mangrove location. Given the confusion with the other terrestrial class, extent produced by COLD is also highly likely to be an overestimate.

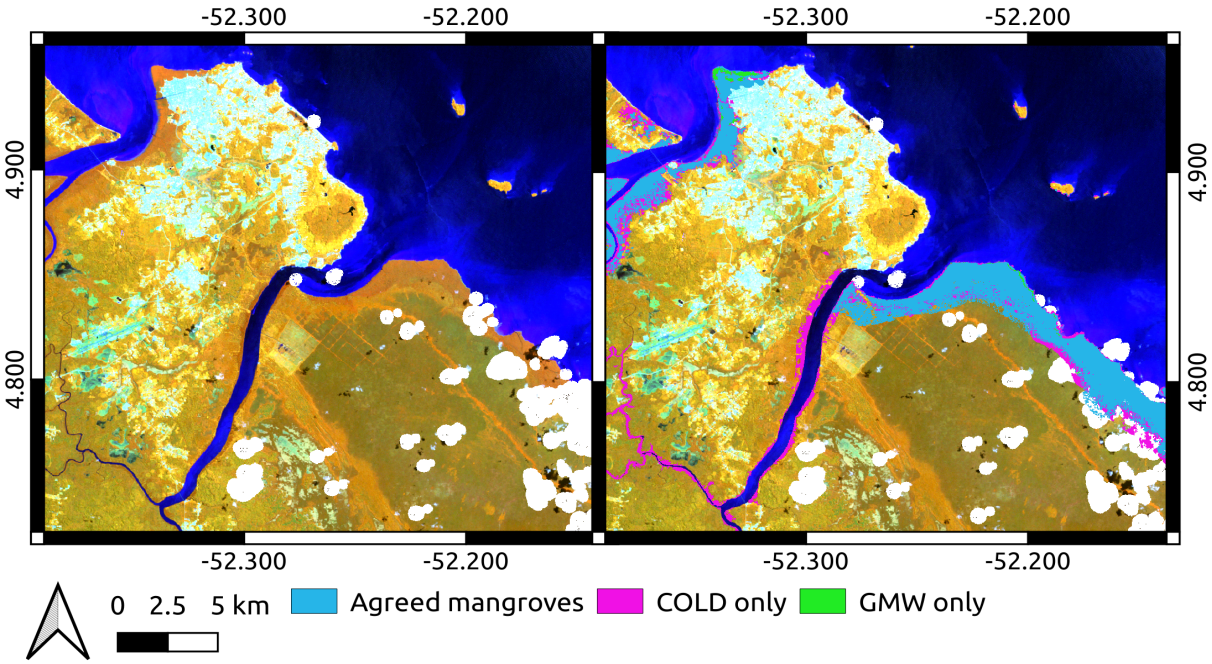


Figure 5.6: Example classification of mangroves for the French Guiana site. Left: false colour Landsat 5 image from July 2010 where Red = NIR, Green = SWIR1, and Blue = Red. Right: mangrove extent as estimated by the two classifiers. Mangroves in this region are spectrally difficult to distinguish from other tropical vegetation.

Years 1990 and 2019 were excluded from extent analysis because more than 40% of pixels for those years had no assigned class. For years 1991-2018, minimum extent recorded was for 1991 (140.3 km^2) and maximum recorded extent was for 2016 (190.3 km^2). Over the the 27-year period, 10 years showed a gain in extent from the previous year and 10 showed a loss, with 7 years registering no change. A gain of around 15% in extent was recorded between 1999 and 2000 (22.2 km^2) (Figure 5.7). As with the Niger Delta site, this gain is likely to be artificial, caused by the launch of Landsat 7 and the subsequent increase in available data. Extent was also identical for the years 1993-1999 inclusive (145.0 km^2), probably due to the lack of data available for the 1990's (Figure 5.3). This explains the absence of any change in extent for that period. Therefore as with the Niger Delta, extent data derived from COLD for this region is not reliable for the 1990's.

5.3.1.3 North Kalimantan, Borneo Island

For North Kalimantan, Borneo the classifier achieved an overall accuracy of 92.3% with a 99% confidence of being between 91.3% and 93.2%. Kappa was the second lowest out of all the sites at 0.82. Quantity disagreement was 0.06 and allocation disagreement was 0.02. As with French Guiana, there was substantial confusion between mangroves and other terrestrial vegetation, with a User's accuracy of 49.9% and a Dice score of 0.64 for the mangrove class (Table 5.6). In particular, there was substantial confusion between mangroves and *nipah* or mangrove palms, which grow in the same lowland coastal areas.

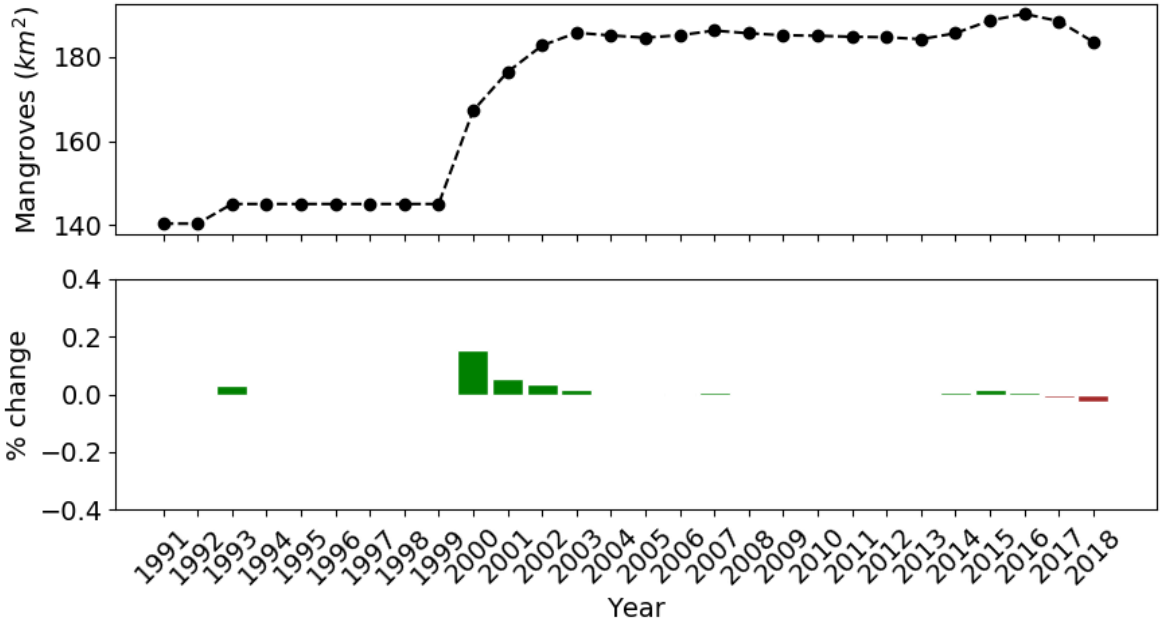


Figure 5.7: Mangrove extent over time for the French Guiana site. Years 1990 and 2019 were excluded from analysis because the class maps for those years contained less than 60% valid pixels.

Table 5.6: Results of spatiotemporal accuracy assessment for North Kalimantan, Borneo island.

		Reference				
		Mangrove	**Water**	**Other**	**Total**	**User's (%)**
Classifier	Mangrove	434	45	390	869	49.9
	Water	46	2892	206	3144	92.0
	Other	11	17	2640	2668	99.0
	Total	491	2954	3236	6681	
	Producer's (%)	88.4	97.9	81.6		**92.3**

When compared to the GMW classification for 2010, COLD estimated mangrove area to be 1035.6 km^2, more than double the GMW estimate of 495.8 km^2. An amount of 55.2% of the area classified as mangroves by COLD was not classified as mangroves by the GMW (Figure 5.8). A total of 6.5% of the area classified as mangroves by the GMW was not classified as mangroves by COLD.

Extent for North Kalimantan exhibited a constant downward trend between 1995 and 2019 (Figure 5.9), with 1995 experiencing the highest mangrove extent (1205.3 km^2) and 2019 experiencing the lowest (952.5 km^2). This is a decline of 252.8 km^2 or 21.0%. Over the 29 years of the study, 20 years recorded a loss of mangrove extent compared with the previous year and 9 years showed a gain. While North Kalimantan did not experience the greatest loss of mangrove in terms of extent or percentage, between 2004 and 2019 it experienced the longest sustained period of mangrove loss out of any of the study sites.

5.3.1.4 Matang Forest Reserve, Malaysia

For the MFR region the classifier achieved an overall accuracy of 97.5% with a 99% confidence of being between 97.0% and 98.0%. Kappa was the second highest of the five sites at 0.96. Quantity disagreement was calculated to be 0.009 and allocation disagreement was 0.02. User's and Producer's accuracies for the mangrove class were

both above 90% (Table 5.7) with a high Dice score of 0.94. Mangroves were mainly confused with other terrestrial land cover; however, in general classification accuracy for this site was very high.

For 2010, COLD estimated mangrove area to be 502.4 km^2, compared to the GMW estimate of 442.1 km^2. 14.5% of the area classified as mangroves by COLD was not classified as mangroves by the GMW. An amount of 2.8% of the area classified as mangroves by the GMW was not classified as mangroves by COLD. Visual inspection suggests that COLD did detect some areas of mangroves which were missed by the GMW classification (Figure 5.10). However, COLD was also slightly more likely than the GMW to overestimate mangrove extent around water bodies.

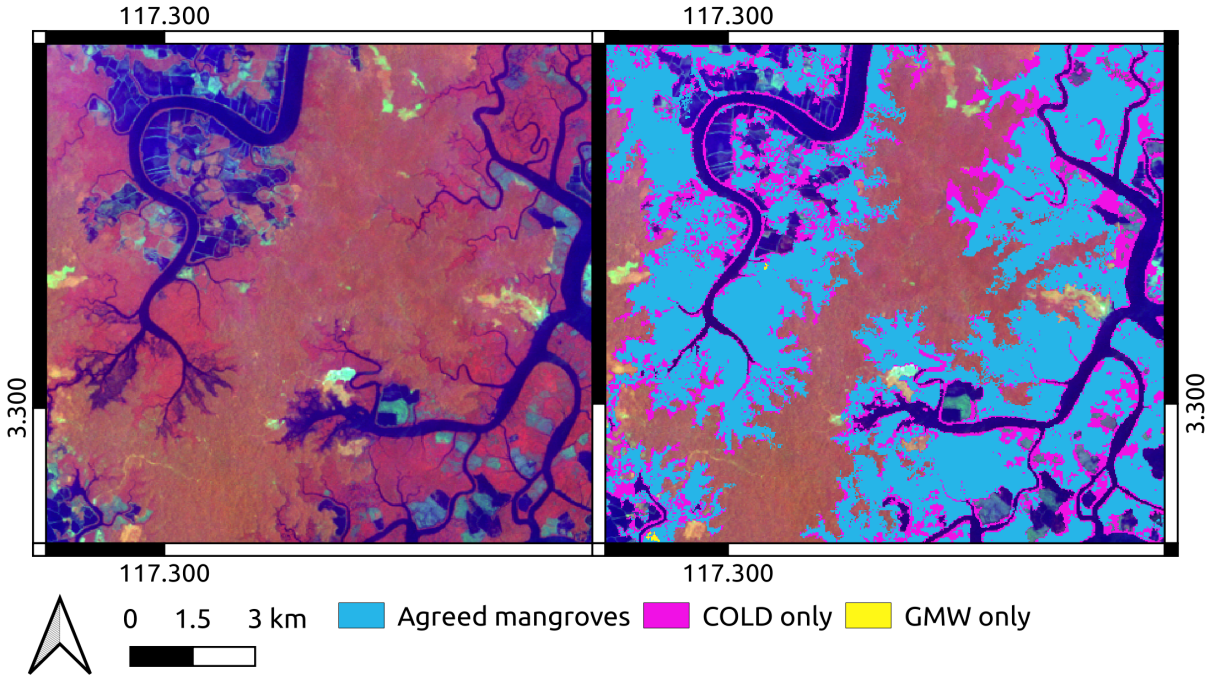


Figure 5.8: Example classification of mangroves in North Kalimantan, Borneo. Left: false colour Landsat 5 image from September 2009 where Red = NIR, Green = SWIR1, and Blue = Red. Right: classification of mangroves showing overestimation of mangroves by the COLD classifier. Mangroves in this region are spectrally difficult to distinguish from other tropical vegetation such as the mangrove palm. There is also some confusion between mangroves and aquaculture ponds.

Table 5.7: Results of spatiotemporal accuracy assessment for the Matang Forest Reserve, Malaysia.

		Reference				
		Mangrove	**Water**	**Other**	**Total**	**User's (%)**
Classifier	Mangrove	1089	37	65	1191	91.4
	Water	19	3035	30	3084	98.4
	Other	22	6	2654	2682	99.0
	Total	1130	3078	2749	6957	
	Producer's (%)	96.4	98.6	96.5		**97.5**

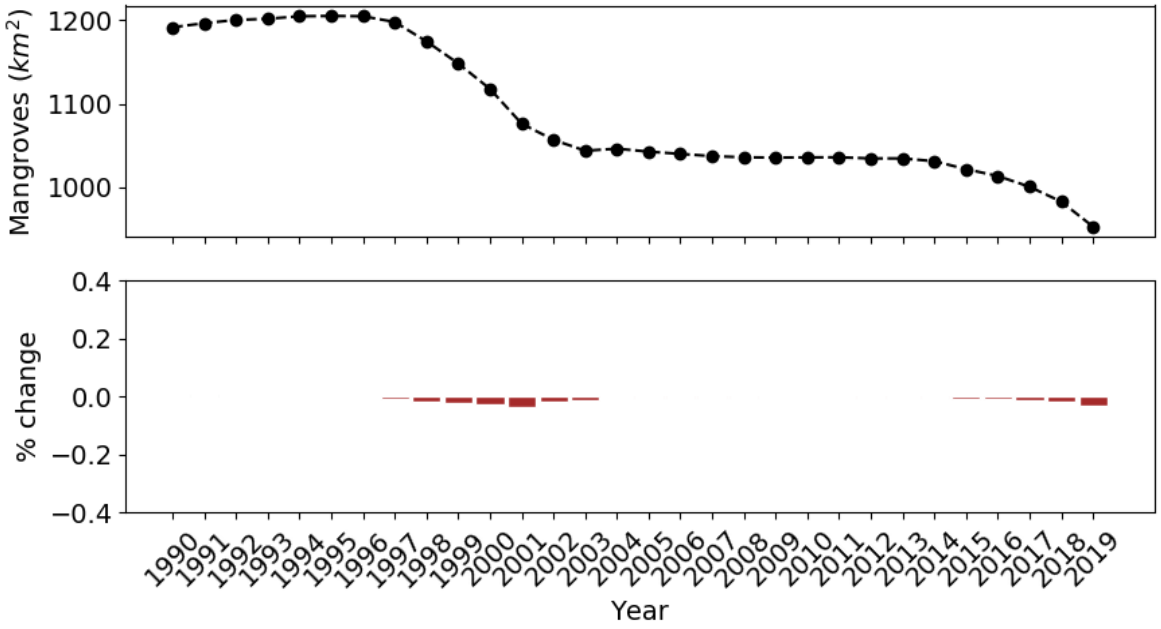


Figure 5.9: Mangrove extent over time for the site in North Kalimantan, Borneo island.

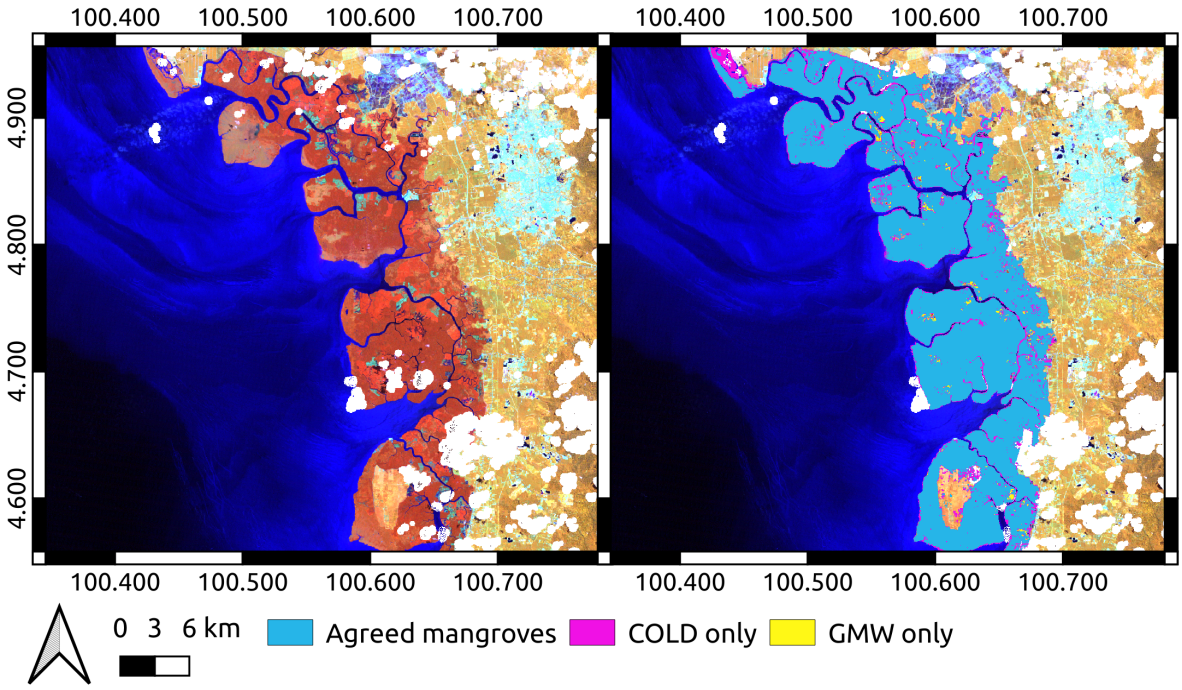


Figure 5.10: Example classification of mangroves in the Matang Forest Reserve, Malaysia. Left: false colour Landsat 5 image from May 2009 where Red = NIR, Green = SWIR1, and Blue = Red. Right: classification showing the high level of agreement between COLD and the GMW for this region. COLD has captured mangroves in the top left of the area which were missed by the GMW.

The maximum extent for the MFR was 516.8 km^2 in 1996 and the minimum extent was 469.6 km^2 in 2019. This represents a reduction of 47.2 km^2 (9.1%) in mangrove extent. A large proportion of that drop occurred between 2012 and 2013, when a decrease of 29.9 km^2 (6.0%) was recorded, although extent did recover somewhat between 2013 and 2014 (Figure 5.11). Over the 29 years of the study, 18 years recorded a loss of mangrove extent and 11 recorded a gain.

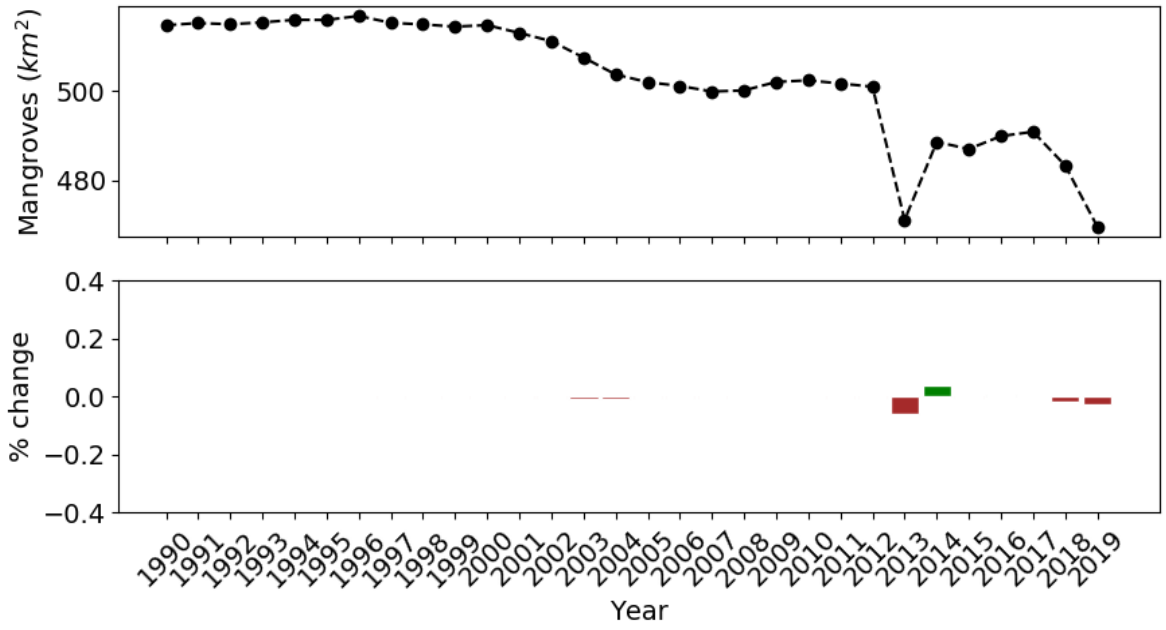


Figure 5.11: Mangrove extent over time for the Matang Forest Reserve, Malaysia.

5.3.1.5 Gulf of Carpentaria, Australia

For the Gulf of Carpentaria the classifier achieved an overall accuracy of 86.1% with a 99% confidence of being between 85.0% and 87.2%. This was the lowest overall accuracy of any site. Kappa was also the lowest out of the five sites at 0.67. While allocation disagreement was in line with the other sites (0.02), quantity disagreement was 0.11, indicating error in the spatial distribution of each land cover class. The majority of this error stems from confusion between the water and other terrestrial classes (Table 5.8). This is likely due to the large tidal and riverine fluctuations in the region, which COLD may have struggled to account for in the modelling process. These changes can also cause problems with manual interpretation of land cover. However, User's and Producer's accuracies for the mangrove class were high (79.5% and 76.0% respectively) with a Dice score of 0.78.

Table 5.8: Results of spatiotemporal accuracy assessment for the Gulf of Carpentaria, Northern Australia.

		Reference				
		Mangrove	**Water**	**Other**	**Total**	**User's (%)**
Classifier	Mangrove	345	7	82	343	79.5
	Water	72	2991	1190	4253	70.3
	Other	37	13	2872	2922	98.3
	Total	454	3011	4144	7609	
	Producer's (%)	76.0	99.3	69.3		**86.1**

For 2010, COLD estimated mangrove area in the Gulf of Carpentaria to be 562.0 km^2, compared to the GMW estimate of 209.1 km^2. A total of 64.0% of the area classified as mangroves by COLD was not classified as mangroves by the GMW. An amount of 3.3% of the area classified as mangroves by the GMW was not classified as mangroves by COLD. Examination of the resulting maps suggests that while COLD probably overestimated mangrove extent, it did capture mangroves missed by the GMW classification (Figure 5.12) and therefore the true extent is likely to be somewhere between the two estimates.

The GOC site experienced more change than any other site over the study period (Figure 5.13). Minimum recorded extent was in 1991 (320.7 km^2) and maximum recorded extent was in 2011 (565.7 km^2). Between 1990 and 2011, the area experienced an overall increase in mangrove extent of 239.3 km^2 (42.3%). However, a substantial drop in extent was recorded between 2014 and 2015 of 154.7 km^2 (31.8%). This drop was almost certainly a result of the 2015 die-back event; however, losses were also recorded for 2011–2012, 2012–2013, and 2013–2014, suggesting that mangrove health may have been declining for several years prior to 2015. Mangrove extent did increase between 2015 and 2018, with extent recovering to closer to 2014 levels, though still substantially less than the maximum recorded in 2011. Over all 29 years, 21 showed a gain in extent and 8 showed a loss.

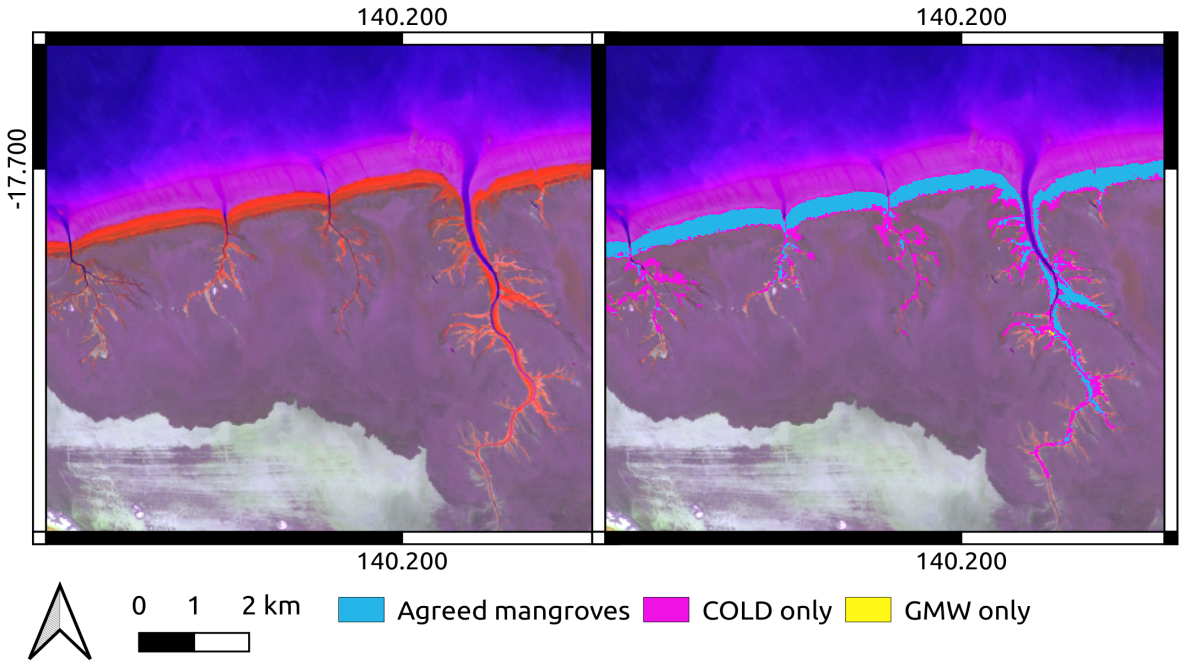


Figure 5.12: Example classification of mangroves in the Gulf of Carpentaria, North Australia. Left: false colour Landsat 5 image from November 2010 where Red = NIR, Green = SWIR1, and Blue = Red. Right: mangrove classification showing areas of mangroves along the coast which were missed by the GMW but captured by COLD.

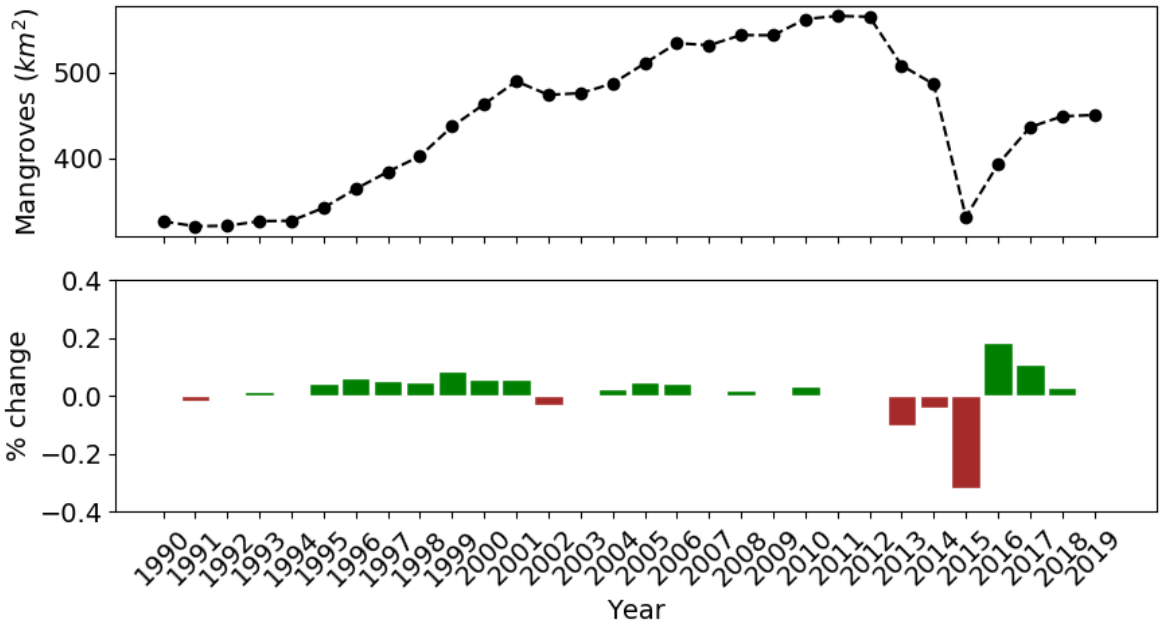


Figure 5.13: Mangrove extent over time for the Gulf of Carpentaria, north Australia.

5.4 Discussion

5.4.1 Niger Delta, Nigeria

Overall classification accuracy for the Niger Delta was very high (98.1%) as were the User's and Producer's accuracies for the mangrove class (93.6% and 97.4%, respectively). COLD therefore proved to be a highly accurate method for mangrove classification in this region. Allocation disagreement was higher than quantity disagreement, indicating more error in the spatial distribution of land cover than in the quantity mapped as each class. However, in both cases the error is very small. The accuracy achieved by the COLD classifier was an improvement over a 2014 study by Kuenzer *et al.* which used Landsat data to monitor land cover changed in the Niger delta between 1986 and 2013 [159]. They reported an overall classification accuracy of 81.6%, however, accuracy for the mangrove classes was only 35-64%. Other studies in this region either do not report accuracy (e.g., [132, 214]) or only look at change rather than specific classes [212]. Previous studies have also been limited to using only single images, often decades apart (e.g., [132, 159, 212, 214]) and no analysis was found using satellite imagery beyond 2013. This study therefore represents substantial improvement over previous studies into mangrove dynamics in the Niger Delta region, in terms of accuracy, data density, and time period covered.

When compared with the GMW map for 2010 for the same region, the COLD method identified around 20% more mangroves by area. The high overall accuracy of the COLD method in this region suggests that the GMW estimate is too low. Other studies place the extent of mangroves in the Niger Delta at around 5000–8600 km^2 [159] whereas this study estimated extent of 2000–3500 km^2. However, this study only covered around half of the area identified as containing mangroves by Kuenzer *et al.* [159] and the GMW (Figure 5.2), meaning that our estimates of extent are within the range of previous studies.

There were issues with data availability for this site. A lack of imagery for much of the 1990's (1991 and 1993–1998) meant that while class maps were produced for 1990–1999, a high proportion of the pixels in those maps were not assigned a class by COLD. It is therefore reasonable to assume mangrove extents calculated for these years would be underestimates. These maps could also not be validated due to the lack of imagery, lowering the total number of validation points available. Given the lack of available data, generating an accurate extent for the Niger Delta over such a long time span is challenging regardless of methodology used. It is clear from this study that, where COLD can fit stable models, it can be used to classify mangrove extent with a high degree of accuracy; however, large gaps in the data record (in this case six years) will lead to an inability to fit stable season-trend models and lead to gaps in classification. Given the length of the gap, once data became available in 1999 it was unlikely to fit the same spectral characteristics as data from six years earlier, meaning that COLD would have struggled to fit a model which covered the intervening years. As a result no model could be fitted to many pixels and therefore no majority class could be determined. This also led to a false increase in mangrove extent being observed between 2000 and 2001, when data availability increased as a result of the launch of Landsat 7 in 1999. The subsequent increase between 2001 and 2003 is also likely to be a straightforward effect of more data being available for more pixels, given that scenes are usually highly contaminated by cloud. The overall dependence on Landsat 7 for this site probably contributed to generally high rates of unclassifiable pixels, given the failure in 2003 of the Landsat 7 Scan Line Corrector (SLC). This lack of available data for some pixels may have lead to underestimation of extent over the 2000's. However, estimated extent for 2010 was higher than reported for the GMW, and no increase in extent was recorded on the launch of Landsat 8 in 2013, suggesting that data availability was not the main driver of change over this period.

Extent was estimated by COLD to be relatively stable throughout the 2000's and 2010's, with a downward trend after 2013 and a drop of 25% between 2018 and 2019. These decreases do not seem to correlate directly with data frequency. Other studies suggest that degradation of the Niger Delta mangroves has occurred since at least the 1980's. Abbas and Fasona [133] and Abbas [132] found a roughly three-fold increase (10.6% to 32.2%) in mangroves and marshland classed as degraded between 1986 and 2008. Omo-Irabor *et al.* [214] found a decrease of 15% in mangrove population between 1987 and 2002 by comparing Landsat imagery between those two years. Kuenzer *et al.* [159] generally found more accretion than erosion between 1987 and 2003, but predominantly erosion from 2003 to 2013. This could account for the increases in extent found by COLD in 2002 and 2003 (Figure 5.5). While our results indicate that mangrove extent did not reduce until the mid-2010's, it is possible that losses inland due to human activity were balanced up until this point by sediment accretion along the coast leading to seaward expansion of mangroves. Overall, our results support the previous evidence that mangrove extent in the Niger Delta is decreasing due to both direct and indirect human activities, and that this decline may have accelerated in recent years.

5.4.2 Area around Cayenne, French Guiana

Classification accuracy for French Guiana was good overall but User's accuracy for the mangrove class was low (61.2%). While COLD was able to accurately distinguish between water and non-mangrove terrestrial land cover, there was substantial confusion between mangroves and other terrestrial vegetation, resulting in overestimation of mangrove extent. Quantity disagreement and allocation disagreement were identical, indicating error in both the quantity and spatial distribution of mangroves. This error appears to be due to the difficulty in distinguishing mangroves from other lowland tropical vegetation. COLD works by distinguishing land cover by its sea-

sonality and while there is evidence that mangrove greenness varies by season [225] that cycle may be similar to that of other tropical forests. While mangroves are generally spectrally distinct, using the NIR, SWIR1, and Red band combination, the boundary between mangrove and other vegetation types can be indistinct where they coexist (Figure 5.6). Two closely related mangrove species, *Rhizophora racemosa* and *Rhizophora mangle*, are known to co-exist with other tropical vegetation in mixed forests further inland where the influx of seawater is highly diluted [79]. Accurate classification within these mixed forest communities will always present a challenge at Landsat scale.

Compared to the GMW extent for the same area, the COLD method classified nearly 30% more pixels as mangroves but did not include nearly 20% of the mangrove pixels identified by the GMW. This indicates substantial disagreement between the maps generated from the two approaches. Given that the GMW classification was used to train the COLD classifier, the lack of overlap between the two classifications is unexpected. It is possible that both classifications are overestimates, and that miss-classification in the GMW training dataset caused confusion. Even with visual inspection, distinguishing the boundary between mangroves and other vegetation in this region is difficult (Figure 5.6). Given the spectral similarity between mangroves and other vegetation, the COLD classifier might perform better with different parameters; for example, a lower value of λ would allow the season-trend models to fit the data more closely. Including non-mangrove tropical forest as an additional class could also improve the classification.

As with the Niger Delta site, data availability for French Guiana over the 1990's was poor, though some Landsat 5 images were available for the 2000's. The availability of Landsat 5 data is probably why more complete maps were generated for French Guiana over the 1990's than for the Niger Delta. While French Guiana also had a six year data gap, Landsat 5 imagery was available for both 1992 and 1999,

meaning more spectral consistency across the data gap. However, while maps were created for French Guiana for 1993-1999 mangrove extent estimated by these maps is identical (Figure 5.7). This is probably because while a baseline stable model could be fitted for many pixels which covered the gap, data on any changes in land cover would not have been available until mid-1999. The jump in extent seen between 1999 and 2000 (Figure 5.7) is therefore likely to be caused by a combination of more pixels being classified overall (due to the launch of Landsat 7) and an actual increase in mangroves, which would have appeared as a more gradual increase had more data been available for the 1990's.

The class map generated for 2019 also had a high quantity of missing data which is difficult to account for. Possibly this is due to high levels of disturbance in the first half of that year. The lower User's accuracy for mangroves in this region means that the data for this site must be interpreted cautiously and minor fluctuations in extent must be discounted as unreliable. There are also few studies which focus on mangrove extent in the region and therefore little data to compare against. However, it is generally known that this is a very dynamic region with ongoing coastal erosion and deposition [10, 79, 87, 89]. Colonisation of new mud banks by mangrove species means that mangrove stands are often small, fragmented, and transitory. For example, Gensac *et al.* found that mangroves could colonise over 90% of a new intertidal mud bank within three years [89] and Gardel and Gratiot [87] found rapid erosion of mangroves around Kourou between 1986 and 2002, with a loss of 60 km^2. While the results generated by COLD are promising, further work is needed to accurately describe and monitor fluctuations in mangrove extent along this coastline.

5.4.3 North Kalimantan, Borneo Island

As with French Guiana, the overall classification accuracy for Borneo was good (92.3%) but User's accuracy for the mangrove class was poor (49.9%), with slightly

more mangroves being miss-classified than were correctly identified (Table 5.6). This was mostly due to confusion with other terrestrial vegetation resulting in substantial overestimation of mangrove extent (Figure 5.8). Quantity disagreement was slightly higher than allocation disagreement, suggesting that in total there was more error in the quantity allocated to each class than in the spatial distribution. This confusion is likely to be caused by mangroves having strong similarities in both distribution and spectral characteristics to other tropical lowland vegetation. In particular, the *nipah* or mangrove palm is very prevalent in this region and, being moderately salt-tolerant, these palms can dominate coastal areas [88]. Allowing closer model fits in addition to introducing mangrove palms as a separate class could improve COLD classification in this region.

For 2010, COLD classified over twice as many pixels as mangroves than the equivalent GMW classification. However, unlike French Guiana, in North Kalimantan COLD did agree with the vast majority of mangrove pixels classified by the GMW. The large difference between the two data sets suggests that as with French Guiana, the classifications produced by COLD for North Kalimantan are unreliable for monitoring small-scale fluctuations in mangrove extent. However, extent generated by COLD did show broad agreement with trends found by previous studies. Langner *et al.* used Moderate Resolution Imaging Spectroradiometer (MODIS) imagery to study forest loss over the whole island between 2002 and 2005 and found a deforestation rate of nearly 8% per year for mangroves, higher than for any other forest type [161]. This was mainly attributed to conversion to crab ponds. A 2016 study by Richards and Friess [253] also highlighted North East Kalimantan as a region with high mangrove loss, suggesting a decrease of more than 10,000 ha between 2000 and 2012. COLD estimated a reduction in mangrove extent of around 27,500 ha (275 km^2) between 1995 and 2019, which is within the same order of magnitude. Richard and Friess did achieve a higher classification accuracy of 71% for the mangrove class, though the

number of validation points was relatively small [253]. In a more general study of forest loss over Borneo, Gaveau *et al.* estimated a 30% reduction in forest between 1973 and 2010 [88].

The overall downward trend reported by COLD for North Kalimantan is therefore likely to be accurate, reflecting the general trend of forest loss over Borneo as a whole. Due to inaccuracies in classification this trend likely includes vegetation with similar distribution and characteristics as mangroves. However, our results still suggest a substantial decrease of 21% in tropical coastal vegetation in North Kalimantan between 1995 and 2019.

5.4.4 Matang Forest Reserve, Malaysia

The COLD method achieved a very high overall classification accuracy for the MFR (97.5%) and a high User's accuracy for the mangrove class (91.4%). This is reasonable when compared with previous studies, which report accuracies of 57-91% depending on mangrove species and age [1, 127, 213, 217]. These figures suggest that COLD can provide an accurate measure of extent over time for the MFR. When compared to the GMW derived extent for 2010 for the same area, COLD produced a higher estimate by about 60 km^2 (442 vs. 502 km^2). Ibharim *et al.* estimated mangrove extent for 2011 to be around 311 km^2, lower than both estimates [127]. Another study estimated mangrove extent in the state of Perak to be nearly 440 km^2 in 2000 [47] compared to the COLD-derived extent of 514.8 km^2 for the same year. The COLD estimate is therefore higher than most other sources. However, results from the current study for the Niger Delta and for previous studies over the Sundarbans mangrove forest [17] and Australia [182] suggest that the GMW often underestimates extent, especially further inland.

The MFR is a closely managed mangrove forest, with mangroves being felled and replanted in decades-long cycles. It was expected that changes in mangrove extent

related to silviculture would be evident in the COLD-derived extent maps. It was found that mangrove extent increased very slightly between 1990 and 1997, before decreasing overall from 1998 onward with a major drop in extent from 2012 to 2013. This roughly corresponds with the results of Ibharim *et al.* who found an overall decrease in mangrove extent of 50 km^2 between 1993 and 2011 [127]. COLD-derived decrease for the same period was 13.7 km^2. Otero *et al.* conducted a detailed study of mangrove clear felling and recovery in the MFR using Landsat data which suggests that the process of clear felling in the MFR is complex [217]. While stands are felled based on a 10-year plan, the process of clear felling is often delayed and can take several years to implement. For example, many sites planned to be clear felled in the early 2000's were not felled until the late to mid-2000's and many sites planned to be felled in 2010–2011 were actually felled between 2012 and 2015 [217]. Sites earmarked for felling in 2017 had not been felled at the time of the study in 2019. This possibly accounts for the drops in extent observed between 2012 and 2013 and to a lesser extent between 2018 and 2019 (Figure 5.11). In addition, Otero *et al.* found that recovery from clear felling events took 5.9 years on average ($\pm$2.7 years) [217].

Given the general downward trend, our results suggest that, while extent was stable until the early-2000's, ongoing delays in the actual dates of clear felling in addition to long recovery times after replanting have resulted in a net loss of mangroves in the MFR over the last two decades. However, it should be highlighted that these changes in extent are still small compared to the overall size of the reserve.

5.4.5 Gulf of Carpentaria, Australia

COLD performed worst on this location in terms of both the Kappa statistic (0.67) and the overall accuracy (86.1%); however, accuracy for the mangrove class was around 80% with the majority of the confusion occurring between the water and other terrestrial classes (Table 5.8). The main cause of this confusion is likely to be

the result of both tidal fluctuations and river flooding, which occurs frequently during the wet season [13]. This means some areas are totally or partially inundated with water for at least some of the year. This causes two problems: firstly, for areas that are inundated on a yearly basis it creates a seasonality which is independent of the underlying land cover type, causing confusion for the COLD classifier, which requires models to fit a period of at least one year. This means that the COLD method is not capable of classifying a region as water for part of the year and land for the remainder. Such areas are likely to be either too variable for a model to be fitted at all, or if the change is consistent enough year to year, it will be accounted for in the model creating essentially a separate class of partially inundated land. Secondly, images for validation were randomly selected, causing potential conflict between the classifier and validation, where areas could be inundated with water during validation which COLD had classified as land. The first problem could be overcome by accounting for these intra-year fluctuations in land cover with a separate class, created by looking for regions with large yearly fluctuations in NDMI. The second issue could be mitigated by using additional validation imagery for areas with fluctuating water levels, i.e., from different seasons and times of day.

For 2010, there was substantial disagreement between our results and the GMW classification. While COLD did overestimate mangrove area to some extent, evidence was also found that the GMW derived extent was an underestimate and missed some large areas of mangroves (Figure 5.12). True mangrove extent in the region therefore lies between the two estimates. Our results broadly agree with the findings of Asbridge *et al.* [13] who found a gradual increase in mangrove extent in the gulf between 1987 and 2014 due to both landward and seaward mangrove expansion. This expansion was attributed to a combination of factors including increased sediment deposition from flood events and increased tidal inundation due to rising sea levels [13]. Asbridge *et al.* observed that of the two most dominant species, *Avicennia marina* is more robust

to change, being better able to tolerate persistent inundation with seawater and more able to withstand storm damage than *Rhizophora stylosa* [13]. This adaptability means that *A. marina* can continue to expand seaward even in areas where dieback of *Rhizophora stylosa* is observed [13].

The drop in extent recorded between 2014 and 2015 was highly likely to be due to the 2015 mangrove die back event, which has been well recorded by other studies [69, 112, 182]. COLD-derived extent estimated a loss of around 32% (185 km^2) during this period, much higher than the 6% (74 km^2) estimate given by Duke *et al.* This could partly be due to differences in methodology. Duke *et al.* estimated mangrove loss by comparing Landsat imagery from April 2015 and March 2016, concluding that the die back event occurred in late 2015. In contrast, this study used majority class in a given year as classified using the models generated by COLD. Given that the drop was recorded as occurring in 2015, this suggests that the models for many pixels registered a change in the first half of 2015 (i.e., that was when new observations longer agreed with the fitted model). Our approach therefore probably captured the decrease as a more gradual event, including some mangroves which were lost in the second half of 2014 and the first half of 2015. This is in agreement with a recent study by Lymburner et al., who found that mangroves in the GOC started to decline in 2014, earlier than previously thought [182]. In addition, the COLD method also likely overestimated mangrove extent in this area, suggesting that it captured a loss of other vegetation as part of the same die back event.

5.4.6 Comparison to the Global Mangrove Watch

In all of the study sites COLD estimated higher mangrove extent for 2010 than the GMW. Between 14.5% (Matang Forest Reserve) and 64.0% (Gulf of Carpentaria) of the area mapped as mangroves by COLD was not classed as mangroves by the GMW. For two of the sites, French Guiana and North Kalimantan, the extent maps

generated by COLD were not accurate enough for the mangrove class for robust comparison to be made. However, for the remaining three sites (in addition to the classification for the Sundarbans produced in Chapter 4) COLD was accurate enough for the over-estimation of mangroves to be higher than the margin of error, meaning that the GMW likely does under-estimate in those regions. Manual examination of the resulting class maps confirmed that the GMW often misses smaller mangrove stands further inland and fringes of mangroves around waterways. This is visible in Figures 5.4 and 5.12. Given that the CCDC/COLD classifiers are trained using the GMW, this raises questions about the accuracy of the GMW dataset and its use in model training.

Based on manual assessment of 53,878 validation points, Bunting *et al.* estimated the overall accuracy of the GMW to be 95.3%. It is important to stress that any global product such as the GMW is likely to have local inaccuracies. While the GMW classifier is manually validated, the scale of application means that comprehensive validation is difficult. The GMW is also itself trained on or otherwise constrained by data drawn from a variety of sources, including optical and radar satellite data, water occurrence maps, shoreline data, and elevation data [39]. The quality of the final classification is then manually assessed and miss-classified areas removed or adjusted where necessary, resulting in a map which is much more accurate than that produced by the original classifier. While this process results in a product that is more accurate overall, it means that there is a substantial amount of human judgement involved in curating the final map. In areas such as North Kalimantan, which are both constantly changing and which consist of a mixture of both mangroves and mangrove-like vegetation, this process is likely to introduce errors which are then propagated through to the COLD classifier. For both North Kalimantan and French Guiana, the GMW is likely to also be over-estimating mangrove extent. While some of this over-estimation is removed during quality assurance, some remains and

this will bias COLD towards over-classification in those areas. In addition, areas of high uncertainty during quality assurance represent a relatively small proportion of the whole mangrove area. This means that when generating a training set for COLD, these difficult areas would have been under-represented compared to the well-represented mangrove signal, and as they are spectrally similar to mangroves COLD classified them as such. The upside of this effect is that COLD also detects mangroves removed from the GMW, because COLD receives little information during training to suggest that fringe areas are not mangroves, and therefore classifies them as being most spectrally similar to the training data for the mangrove class.

It is also important to note that no classifier can guarantee complete accuracy. When considering approaches which provide yearly class maps, there will always be some error or disagreement because of underlying land cover change. For COLD, a cutoff was chosen whereby if a pixel was classified as mangroves for more than 50% of the year based on the fitted model, the pixel was classed as mangroves for that year. This already introduces a substantial margin for disagreement which is difficult to resolve, and the significance of which depends largely on the specific application.

As the GMW product continues to be used for both global and local applications, errors of both under- and over-estimation will be identified and corrected in new versions. Despite local variations in accuracy, the GMW still currently represents the best estimate of global mangrove extent, and therefore the most globally consistent dataset for further work on mangrove classification. Currently, the GMW does not attempt to classify vegetation such as mangrove palms which is spectrally very similar to mangroves. Future work could focus on using the GMW as a baseline from which additional classes can be added, especially for challenging areas. This would increase the accuracy of downstream classifiers such as COLD, and further reduce the need for manual quality assurance.

5.4.7 Efficacy of the COLD Algorithm for Global Mangrove Monitoring

While many previous studies have utilised EO data for mangrove monitoring, few have attempted to apply a consistent methodology over a large and diverse spatiotemporal extent. Given the vulnerability of mangroves to the effects of climate change, understanding historic changes to the global mangrove population is vital to tracking and protecting mangroves as an important global resource. Our results indicate that the COLD approach is a promising methodology for solving this problem. In particular, the COLD method resulted in highly accurate maps of mangrove extent for the Niger Delta from the early 2000's onwards despite data being limited in the region due to cloud cover. The method also achieved good results in both the MFR and Gulf of Carpentaria sites, detecting changes in extent that could be related to external factors such as silviculture practices in Malaysia and the 2015 mangrove die back event in Northern Australia. In all three of these sites there was evidence that COLD was able to detect areas of mangroves missed by the GMW, suggesting that the GMW estimates are generally low. In particular, our study found that the GMW may have underestimated mangrove extent in the Niger Delta by nearly a fifth for 2010. Our results suggest that where there is an existing, highly accurate dataset such as the GMW, COLD can be used for temporal extrapolation, reducing the need to repeat the original methodology. The overall accuracy of 92.7% with a User's accuracy of 77% for the mangrove class and a Dice score of 0.84 indicates a reasonable level of agreement between predicted and actual land cover.

The present study also represents an increase in feasibility over a previous study [17], whereby an increase in the time between model updates from one to 90 days substantially decreased processing time from around two days per Landsat footprint to less than a day. While this may increase the number of false land cover changes detected [344], the trade-off is reasonable, especially for broader detection of land

cover where accurate detection of the individual dates of change is less vital. The period between updates could be increased to further decrease processing time for very large scale land cover detection, or when working on systems with limited computing power. Zhu *et al.* suggest that updating the model for every available observation produces the best results, but found updating every year (365.25 days) to be a reasonable compromise, while updating only every two or three years substantially reduced change detection accuracy [344]. Once COLD has been run over the historical archive, there is also potential for it to be used for ongoing monitoring. New scenes can be compared to the existing models and pixels flagged as no change or potential change on a near real-time basis [340, 344].

Accuracy was lower for the French Guiana and North Kalimantan sites, primarily due to confusion with other tropical vegetation. Classification accuracy for these sites could be improved with the introduction of more land cover classes. However, this presents a difficulty in areas where mangroves and other ecologically similar vegetation co-exist without clear boundaries. At Landsat resolution of 30m, distinction between these land cover types may not be possible without more manual involvement during both the generation of training data and of the final land cover maps. A combination of introducing specific additional classes, such as one for mangrove palms, and tweaking the closeness of the allowed model fit, could improve classification in these regions and is worth further investigation. A possible approach would be to take sample pixels which were manually verified as being dominated by each specific class, then varying the fit of the models to investigate at what point they produce outputs discernibly different to a classifier. If distinction by seasonality is not possible then COLD will likely always produce overestimates for some regions. However, even in these cases the COLD method has utility for more general monitoring of tropical coastal vegetation.

5.5 Conclusions

Awty-Carroll *et al.* [17] demonstrated that mangroves could be mapped through the Landsat time series using a model-based approach, such as the COLD algorithm [344]. Therefore, this study aimed to identify whether the approach was transferable and in the future could be applied on a global basis to map historical mangrove extent. The primary concern was the availability of Landsat imagery, where in regions such as the Niger Delta, the availability of historical data is limited. For these regions, mapping using the COLD approach is only possible once Landsat 7 data become available (i.e., ∼2000). However, for many regions of the world, where Landsat 5 was more widely acquired, mapping back to 1990 can be reliably achieved. The second consideration was compute time. Model-based approaches require significant computation time and to undertake a global analysis of approximately 1800 Landsat row/paths could be computationally prohibitive. While further compromises in model accuracy could be made to reduce the computation time, this study has demonstrated that each Landsat scene could be processed in under a day using approximately 600 cores, depending on how many images are within the time series. Therefore, the application of the COLD approach to mapping historical mangrove extent globally is considered viable, providing high quality mapping summarised on an annual basis while also accounting for seasonal changes. Additionally, reflectance trends can also be retrieved, allowing for the identification of degradation (e.g., [17]) and COLD also has potential for near real-time mapping and alert systems. However, if the objective was to make the earliest map possible from the Landsat archive, then the 'spin up' period required for the COLD algorithm is prohibitive and either a scene by scene approach or map-to-image based change approach as used in Thomas *et al.* [285] would be appropriate.

Chapter 6

Conclusions

This section will summarize the main conclusions of the thesis in terms of the aims and objectives. Briefly, these were to document and test existing methods for land cover/land use change monitoring with remote sensing time series, and to investigate the feasibility of these methods for large-scale analysis of change over large spatiotemporal scales. In Chapter 2, a review of the current approaches to land cover change monitoring is presented. Chapter 3 compares four of these methods using simulated NDVI time series to represent a wide range of change types, noise, and levels of missing data. Chapter 4 applies one of these methods, CCDC, to a specific use case of monitoring mangrove change in the Sundarbans region including an assessment of damage from a major destructive event, Cyclone Sidr. Chapter 5 applies an updated form of the same method, now known as COLD, to long term monitoring of five sites across the globe with know drivers of mangrove change.

6.1 Usefulness of simulations in remote sensing

Time series simulations have not widely been used in remote sensing studies. Where studies have employed simulations, use has been limited in terms of the number of methods being evaluated (e.g. Verbesselt *et al.* [296]) or the type of change represented (e.g. Forkel *et al.* [75]). This is at least partly because in remote sensing there is a vast amount of free data available for time series construction, and gener-

ation of additional data is seen as unnecessary. Also, simulations are often seen as inferior to "real" data, and signals derived from RS sources are complex to simulate in many ways. Rarely is there a clean signal, with observations being contaminated with sensor and atmospheric noise. RS time series are often incomplete, with missing data that can occur randomly (as in the case of clouds) or in a known pattern (as in the case of the LS7 SLC failure). Many studies have been devoted to cleaning these signals to generate useful time series with values representative of the land's surface. It can be argued that simulated time series will always struggle to represent the variation present in true data.

However, simulations are used across many areas of science to simplify problems and create known scenarios for testing. The problem with using real time series is twofold: Firstly, it is difficult to find a representative sample of both underlying land cover signal and change type; and second, it is difficult to accurately judge when a change has occurred and what the nature of that change was. This is a general issue in RS, where true ground truth data is hard to come by. While it might be possible to collect data for specific scenarios (i.e. forest monitoring where inventory and felling data is available), obtaining ground truths for changes over most of the Earth's surface is not possible. Even where ground truth data does exist, the data itself will be limited to specific types of change (e.g. abrupt change due to fire or clear felling). Changes in phenology, for example, are very difficult to track in satellite imagery except at very broad scales, and while a lot of ground truth data is available (e.g. for crop growth or bud burst [20]) this is often collected at scales too small to be useful in RS studies. Given this, there is a clear gap in the research that simulations can fill. Simulated time series can contain known phenological parameters, known timings and types of change, and known quantities of noise and missing observations. While it can be argued that simulations are too perfect and cannot represent the true randomness of RS observations taken over time, it can also be argued that if an

algorithm cannot perform well on this type of "perfect" data, it cannot be expected to perform well on real data.

In Chapter 2, four related algorithms were tested using simulated data. All of these algorithms made assumptions about the data being processed, e.g. that the signal to noise ratio was high enough for the signal to be interpreted and that each time series did have periods of stability where phenology and trend remained similar enough to be represented by the same model over multiple years. Part of the purpose of that study was to identify the points where these assumptions broke down. Interestingly, the study showed that while all of the methods used a very similar form of linear regression to model the time series, the method of break detection used did have a substantial impact on robustness to noise and missing data. BFAST, the only method to operate over whole time series rather than detecting change as deviation from a stable history period, performed very consistently across all noise levels. This suggests that one limitation of the near-real time methodologies is that they have no ability to know the true distribution of the signal, and their performance will depend strongly on how well the history period represents that distribution. Methods such as CCDC/COLD, which adapt to small changes in the signal over time, risk higher error rates if the fitted model is not updated frequently enough, because new observations become less and less likely to fit the original model.

Chapter 2 also showed that even when presented with simulations with clear and easily identifiable changes in phenology (length of season, number of seasons, and amplitude of signal), all of the methods struggled to correctly identify a change. Sometimes, these changes were instead registered as abrupt or trend changes. This suggests that this type of modelling struggles to properly capture seasonality. Generally, the studies presenting these methods did not focus on testing their capacity to detect phenological changes [35, 297, 342], with the exception of BFAST [296]. It could certainly be argued that they were not designed to be able to detect these

types of change. However, accurate capture of the underlying phenology is important to change detection, especially in long-term time series where climate change effects may be present. By using simulations, the ability to detect and/or incorporate such changes can be assessed, and future work can better take account of this limitation where it exists.

Because simulations have been so little used, it is difficult to properly ascertain their place in RS research. Testing with simulations clearly offers benefits, and as applications broaden to change detection on continental or even global scales, the use of simulations for training and testing LULC change algorithms may increase. Additionally, new methods may emerge that make it easier to create realistic simulations. For example, Generative Adversarial Networks (GANs) are a form of neural network architecture with two models: A generator and a discriminator [57]. The generator outputs simulated data based on a range of input data from which it learns the data distribution. The discriminator is trained to discriminate between generated and real data. The generator is told which outputs are seen as real and which are identified as fakes. Over time, the generator creates more realistic outputs until it becomes very difficult for the discriminator to distinguish fakes from real samples. Such methods could help in generating time series which more accurately represent the distribution and variability of real RS time series.

6.2 Time series modelling for change detection and interpretation

In Chapter 2, a wide range of methods were discussed for monitoring of LULC change in RS time series. Such methods present advantages over more traditional approaches involving image-to-image comparisons. Namely, they reduce the reliance on whole cloud-free images or composites (ideally from the same season to avoid phenological change effects), reduce the influence of noise and outliers by taking temporal context

into account, and allow for easy interpolation and prediction of values. In dynamic regions such as mangrove ecosystems where ephemeral land cover change is constant, extracting true and meaningful change information can be difficult when the wider temporal variability in observations is not taken into account. However, time series modelling methods present their own challenges. They can require extensive fine-tuning, with multiple parameters which need to be adjusted for each specific use-case. Methods can be difficult to implement if code is not readily available, and in the case of near real-time processes, require identification of a stable history period from which change can be detected as deviance. In addition, dense time series modelling over large spatiotemporal scales is computationally expensive, limiting uptake to smaller scale studies or those with access to large scale computing facilities and the expertise to use it. Choosing which method to use is not always straightforward, and can have a significant impact on results; a 2017 study by Cohen *et al.* found that change maps derived from different algorithms rarely agreed about either the location or the timing of changes [50].

The results presented in Chapter 3 of this thesis suggest that methods may also have limitations that are not immediately apparent without systematic testing. For example, a gap appears to exist between methods aiming to analyse phenology specifically (such as the approaches included with TIMESAT) and inter-year modelling approaches such as BFAST and COLD. The former are capable of extracting a range of detailed metrics, such as start of season and peak greenness, which are useful representation of the underlying phenology and can be used to monitor phenological change between years. The latter tend to make fairly rigid assumptions about seasonality, using season-trend models which contain seasonal coefficients but allow for little variation year to year. This broad stroke approach is adequate for many applications, but as shown in Chapter 3, it makes these models poor choices for any use case involving the detection of phenological change. Another consequence is that

because these models aim for the best fit over multiple years, the fitted model will likely underfit on some years and overfit on others, reducing change detection efficacy. It is important to note that this does not mean these methods are wholly ineffective; simply that their limitations must be acknowledged. In the case of CCDC/COLD, the algorithm is designed for monitoring of class change rather than phenological or trend change, and no claim is made otherwise.

There are also broader issues within the field of change detection. Since the opening of the Landsat archive in 2008, researchers have had access to a huge amount of historical satellite data. As a result, studies have increasingly employed time series approaches that can make use of long satellite image time series [338]. However, this presents an issue with validation of change events and land cover maps. No other sensor exists which matches the spatiotemporal extent of the Landsat missions, and gathering actual ground truth data (e.g. forest inventories) is even more difficult in a historical context. In many cases, determining the actual timing and nature of change events relies on human interpretation (e.g. [76, 211, 323]) which is even more difficult when alternative data sets are unavailable for the study period. The lack of ability to validate outputs with an independent data set is a substantial limitation that can only partly be overcome, by using independent data where possible (e.g. data from Sentinel-2 or from airborne sensors) and by expert interpretation. Linking change drivers (such as climate change) to observed land cover changes is not straightforward; correlation does not mean causation. In Chapter 4, a drop in mangrove extent was linked to a major cyclonic event. This event happened recently enough (2007) that other satellite data was available along with derived estimates of mangrove loss which could be compared against. Even in this case, the results must be interpreted cautiously, and the aim was to produce a robust low-end estimate of loss rather than a maximum. In Chapter 6, while broad changes in mangrove forest extent could be identified along with possible change drivers, the spatiotemporal extent of the study

along with lower classification accuracies meant that reliable interpretation of change events was impossible for the Borneo and French Guiana sites. Care must always be taken when considering change estimates that lie within the margin of error. However, there is no doubt that determining long-term trend changes in satellite image time series is extremely valuable, particularly in tracking the effects of climate change.

While the Landsat archive remains a valuable resource, there is no doubt that an increase in available satellite sensors along with the increasing availability of airborne imagery and high performance computing platforms (e.g. JASMIN) means that shorter term analysis of dense time series is becoming easier, more accessible, and more reliable. In a recent paper, Woodcock *et al.* suggest there has been a paradigm shift in remote sensing away from change detection and towards LULC monitoring [316]. While the difference between the two may seem subtle, useful and accurate real-time monitoring (e.g. for monitoring forest fires or earthquakes) will rely on reliable operational systems which utilise dense time series data from a variety of sources. CubeSat satellites such as those used by Planet (Planet Labs Inc., California) [234] can now produce daily global imagery at 5 m spatial resolution, representing a substantial advance in the field of LULC monitoring. As the pool of potential data sets grows, so does the potential for data fusion methods which further increase temporal resolution and for methods which facilitate the interpolation of higher temporal resolution imagery with higher spatial resolution imagery (e.g. Prediction of Sentinel-2 imagery from Sentinel-3). With daily or near-daily observations, the efficacy of current modelling methods will be further tested and new methods will undoubtedly be developed to take advantage of higher data availability for more accurate LULC classification and monitoring.

6.3 Feasibility of global-scale monitoring with remote sensing time series

Use of moderate resolution satellite imagery for operational global scale monitoring has been identified as an area of continued growth [316]. Access to high-specification local computing systems, high performance computing platforms, and services such as AWS and GEE has undoubtedly increased the ability of researchers to carry out LULC change studies on large spatial as well as temporal scales. Currently, while many studies have used lower spatial resolution imagery such as MODIS (Max. 250 m resolution) or AVHRR (max. 1.1 km resolution) to monitor change on country, continental, and even global scales (e.g. [4, 140, 161, 205, 308]), far fewer studies have attempted this with Landsat or Sentinel imagery (10-30 m resolution). However, there are examples of operational large-scale land cover monitoring products, such as those produced by the Global Forest Watch [318] and Digital Earth Australia (DEA) [179]. Both of these efforts have produced multi-year products at Landsat spatial resolution. Despite these examples, a number of barriers remain to researchers looking to monitoring LULC change on large scales, including the need for significant processing power and data storage, the need for large-scale validation data sets (which may require international co-operation to obtain), and the need for substantial manual quality assurance and interpretation. The requirement for robust accuracy assessment and QA is particularly important for operational products with a range of end users including policy makers.

The question of methodology also remains. The Global Forest Watch (and by extension the Global Mangrove Watch) products focus on a narrow range of land cover classes. The DEA maps use the United Nations Food and Agriculture Organisation (FAO) Land Cover Classification System (LCCS) to classify a range of land cover types in a hierarchical system, but this is of course limited to Australia. To investigate the feasibility of a more automated approach based on time series modelling,

a potential new methodology for mapping global mangrove populations was implemented and discussed in Chapter 5. Based on COLD, this method was applied to five sites across the globe representing a range of mangrove species and change drivers. While this study was successful in classifying mangroves with high accuracy at three sites, it also highlighted some of the challenges faced when implementing LULC monitoring across diverse regions, even when considering a very limited number of land cover classes. For example, the method performed very well in the Niger Delta and Malaysian sites, where the boundary between mangroves and other land cover classes was very distinct. However, it struggled in the French Guiana and North Kalimantan study sites due to the presence of spectrally similar tropical vegetation. Classification was also less accurate in the Gulf of Carpentaria, where ephemeral land cover change due to frequent floods and tidal effects likely caused some confusion both in model fitting and accuracy assessment. The huge array of different ecosystems, species, land cover types, and change drivers means that adaptation of methodologies may always be a necessary step in creating global LULC products. A process which works well for one region may not work well for another. This need for parameter tuning and/or methodological adaptation creates a requirement for manual involvement. Therefore, researchers must accept the requirement for a substantial amount of time to be spent either on method tuning or on manual QA, in addition to the time spent on hand-labelling training data.

The amount of human input required for this type of mapping means that researchers are less likely than commercial enterprises to have the time and funding available for large scale mapping projects. A recent effort by Esri (the company behind ArcGIS, a popular platform for GIS users) resulted in the release of a global 10 m resolution land cover map for 2020 [72]. Based on Sentinel-2 imagery, the map was trained on 5 billion hand-labeled pixels taken from over 20,000 sites [72] and achieved an overall accuracy of 86% across 10 land cover classes (though accuracy for some

classes was substantially lower). Generated using a deep learning model, this map represents the cutting edge of global land cover mapping, and with a trained classifier, production of maps for other years should be possible. However, the approach used still relied on the use of composite imagery and is still only allows for monitoring of changes in land cover class. As shown in Chapter 4, time series method have the potential for the generation of much richer datasets including estimates of condition change in addition to changes in extent. Time series methods are also potentially a much more robust method for LULC class estimation because they take into account the underlying phenological cycle of vegetation. The COLD method is already being considered by the USGS-NASA team for monitoring of the entire conterminous United States [344], and it is likely that in the coming years we will see more efforts involving large scale applications of time series modelling approaches. In addition, advances in deep learning such as the development of Convolutional Long Short-Term Memory networks [274] which are specifically designed for classification and prediction of spatiotemporal datasets means we are likely to see continued advances in both the efficacy and efficiency of LULC change algorithms.

6.4 Main findings

A set of research questions were posed in Chapter 2 which the thesis has aimed to investigate. The main findings of the thesis are summarized here.

- The use of simulated NDVI time series to assess the efficacy of time series modelling approaches for change detection was highly informative. While simulated time series cannot replicate the complexity of real remote sensing data, they can provide vital insight into the ability of approaches from the literature to accurately detect change. The simulation methodology presented in this thesis allows for evaluation of each method against a wide range of change scenarios and allows for comparison between methods based on temporal accuracy and

rates of commission and omission error. Chapter 3 provides recommendations as to the applications of each method. The full set of simulated time series is available online [16].

- EWMACD was found to be the most effective method for detection of small magnitude changes and changes in phenology (e.g. number of seasonal peaks or increase in signal amplitude). CCDC (now known as COLD) performed well at detecting larger magnitude changes and tended to ignore smaller magnitude changes and phenological changes. CCDC also performed well in cases where missing data levels were high. BFAST failed to perform well at detecting phenological changes despite this ability being a key feature of the original paper [296]. BFAST Monitor performed poorly, frequently missing true breaks while also having a high false detection rate.

- Given its robustness to missing data and ability to detect large magnitude changes (i.e. changes in land cover) while ignoring smaller magnitude changes (e.g. as caused by ephemeral change), CCDC/COLD is an appropriate method for monitoring extent change in mangrove ecosystems.

- In Chapter 4, the benefits of applying the CCDC method to a use case of mangrove forest monitoring in the Sundarbans were demonstrated. While computationally expensive, CCDC was able to produce highly accurate maps of mangrove extent in addition to providing information on cyclone damage and long-term recovery trends. There is evidence that the CCDC method was able to detect small stands of mangroves missed by the Global Mangrove Watch dataset. Chapter 4 presents the most comprehensive assessment of mangrove forest extent in the Sundarbans so far, with previous studies focusing only on smaller geographic or spatial scales.

- The results of the COLD method showed that with the exception of forest loss caused by Cyclone Sidr, the extent of the Sundarbans mangrove forest remained roughly stable from 1990 until around 2011, when the population starts to decline. This downward trend is concerning, and indicates that further monitoring is needed to establish whether losses will continue or whether extent will stabilize and/or recover.

- Chapter 4 found evidence that although mangrove extent recovered quickly after an extreme cyclone impact, the effect of the cyclone on mangrove health persisted long-term. When looking at trends in NDVI derived from CCDC, around 50% of the damaged area had not recovered to pre-cyclone NDVI levels by mid-2018. There is also evidence of a long-term decline in the health of *H. fomes* trees in the region, probably due to changes in water salinity caused by rising sea levels.

- In Chapter 5, the COLD method was applied to five study sites representing a range of mangrove species and change drivers. COLD was demonstrated to be highly effective for classification in three of the five sites, and accurate yearly estimates of mangrove extent were produced. In two sites, North Kalimantan and French Guiana, COLD struggled to delineate mangroves from other tropical vegetation. This limitation will be important to address in future studies. However, for North Kalimantan in Borneo the results from COLD did suggest a general downward trend in tropical vegetation extent, which correlates with other studies into vegetation loss in the region.

- COLD was successfully applied to the Niger Delta mangrove forest, generating extent maps from 2000 to 2019. Currently this is the most comprehensive record of mangrove extent available in the Niger Delta, which is under-studied due to the lack of available satellite data (the result of high cloud cover). Mangrove

extent for the Niger Delta was found to be relatively stable for much of the 2000s, but a decrease was observed during the 2010s.

- COLD-derived extent estimates for the Matang Mangrove Forest Reserve in Malaysia showed an overall downward trend. This result is contrary to previous studies, which show extent in the reserve as remaining relatively stable. However, due to the long-term nature of forest management in the region, this trend could be reversed in the future and may be caused by delayed harvesting of mangrove stands. These results suggest that further monitoring is needed in the region to establish whether the decrease in extent is permanent.

- Results in Chapter 5 for the Gulf of Carpentaria reveal that mangrove extent began to drop as early as 2012-2013, suggesting that the 2015 dieback event actually began much earlier than previously suggested in the literature. Prior to 2011, mangrove extent in the area was steadily increasing. While this drop cannot be explicitly linked to the major loss of mangroves experienced by the region in 2015, it suggests that mangrove health was already declining when the dieback event occurred. The results also show some recovery of extent from 2016 onwards, suggesting that mangroves are re-establishing in the region.

- When compared with the GMW extent baseline for 2010, CCDC/COLD gave higher extent estimates for all six study sites. However, for two sites (North Kalimantan and French Guiana), accuracy for the COLD classifications was too low for extent to be considered reliable. For the remaining four sites, between 10.7% and 64.0% of the area classified as mangroves in 2010 by COLD was not classed as mangroves by the GMW. Manual interpretation suggests that while there may be some over-estimation of extent by COLD, the GMW does tend to underestimate mangrove extent. This is significant given that the GMW often fails to include smaller, more vulnerable mangrove stands. The discrepancy is

likely introduced during post-processing, when many smaller stands are probably removed from the GMW maps during manual Quality Assurance. This results suggests that while the GMW may be a robust minimum estimate of global mangrove extent, true global extent could be substantially higher.

- Overall, the results presented here show a complex picture of global mangrove dynamics. Several sites (The Sundarbans, Niger Delta, and Matang Forest Reserve) show a possible downward trajectory of mangrove extent in recent years that must be monitored going forwards. For the Sundarbans and Gulf of Carpentaria sites, which both experienced catastrophic loss of mangroves, there is evidence of mangrove resilience as populations recover. However, there is also evidence of long-term degradation of mangrove health in the Sundarbans as a result of climate change and the impact of cyclone Sidr. These findings contribute to a growing body of evidence that mangrove forests are under threat from the cumulative impacts of climate change.

- This study has demonstrated the feasibility of applying a dense time series change detection and classification approach to remote sensing data. While computationally expensive, the results show that CCDC/COLD can provide much more detailed information on mangrove health than classification-only approaches. CCDC/COLD also allows for the generation of yearly extent maps which take complex intra-year and inter-year mangrove dynamics into account, meaning that mangrove extent patterns can be examined in detail over decades. A Python version of COLD has been implemented and published online for use in future studies [15].

6.5 Future work

In Chapters 2-5, a range of methodologies for change detection in remote sensing time series were described, investigated, and applied to a set of real use-cases. The results discussed in Chapter 3 represent the first known study to utilise time series simulations for rigorous analysis of such approaches. However, the study is limited in that the simulated data cannot be said to truly represent the full range and distribution of variability present in satellite image time series. Such time series are subject to a variety of outside influences such as atmosphere, clouds, sensor defects, and pre-processing errors. Improving the generation of such simulations to better represent reality would increase their usefulness, with potential to create a benchmark data set against which more algorithms, both new and existing, could be tested. While the field is not lacking in data sources, there remains a need for a robust and objective test bed. This could be achieved through better understanding of the variability in RS time series, including potential use of deep learning approaches such as Generative Adversarial Networks. In addition to the generation of appropriate simulated benchmark datasets, there is also a need for a comprehensive global training and validation data set for assessment of land cover monitoring algorithms. An open database of such samples would be extremely valuable to the remote sensing community, and allow for much easier comparison between algorithms.

In terms of application, the CCDC/COLD method showed promise for monitoring both mangrove extend and condition over large spatiotemporal scales. The ability to extract information in condition as well as extent is important, especially given the time and resources needed to run such algorithms. More detailed work is needed to assess how well these algorithms perform across different land cover types and what information can be extracted. For example, the ability of CCDC/COLD to distinguish between stable, degrading, and improving mangrove communities discussed in

Chapter 4 could be investigated and applied to other vegetation types. Improving the speed of time series analysis when applied to long time series will also be vital in increasing uptake and operational viability. In recent years, access to Graphical Processing Unit (GPU)-accelerated libraries has greatly improved. These libraries can make use of the parallel processing capabilities of GPUs to speed up machine learning algorithms by tens to thousands of times. Libraries such as Rapids [245] implement drop-in replacements for many Python functions, making it very easy to move computation from the CPU to the GPU. Re-engineering of current algorithms to make use of such libraries could vastly improve computational performance. Finally, there is also the potential to combine pixel and object based approaches for time series analysis. For example, many regions are effectively homogeneous over large areas, such as forests or deserts. In many cases it does not necessarily make sense to run analysis on a per-pixel basis. Instead, stable areas of land cover could be identified and processed as objects using summary values, with region-growing techniques used to identify edges or areas of change which need to be processed in more detail.

The rise of deep learning approaches has been exponential in recent years and this will undoubtedly continue into the near future. Currently applications using remote sensing time series have been limited. However, approaches such as Long Short-Term memory networks and more recently self-attention architectures [294] are specifically designed for time series processing and prediction. Recent studies have begun to demonstrate the potential of such approaches when applied to satellite image time series [175, 183, 222]. In particular, the ability of neural network architectures to recognise spatial as well as temporal patterns is likely to be highly useful in land cover identification, allowing for spatial relationships and textures to be combined with phenological information. However, it remains to be seen whether deep learning models can extract the same level of detail in terms of change magnitude and type, especially when considering condition change. These algorithms are often referred

to as “black boxes” because the method by which the model arrives at an output is difficult to interpret. Regardless of this, deep learning will undoubtedly become more popular within remote sensing due to the ability of deep learning models to capture complex non-linear relationships difficult to represent with traditional approaches.

Appendix A

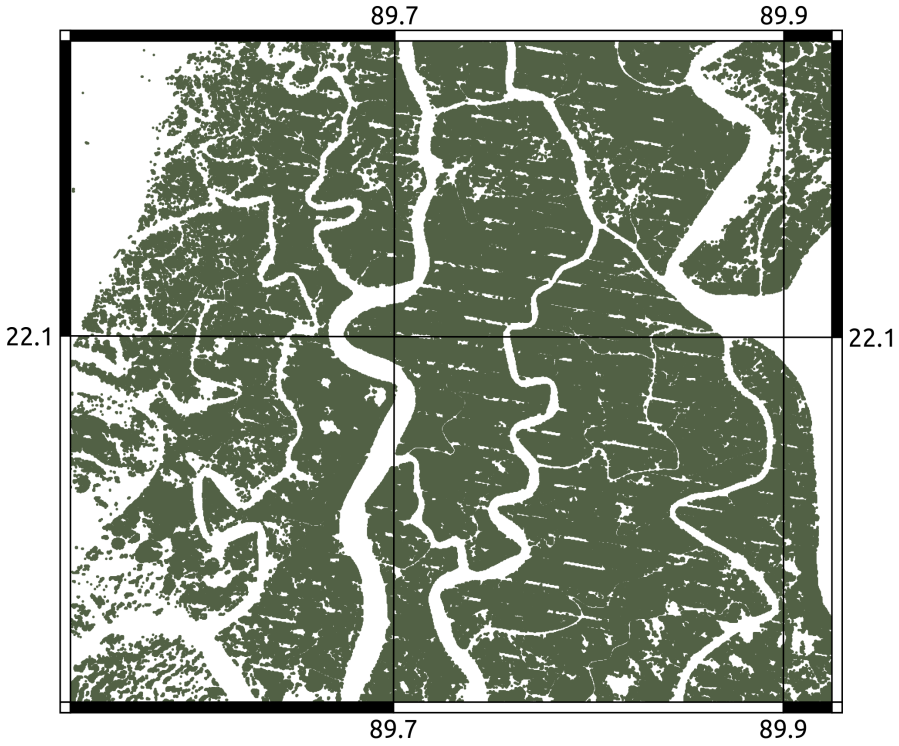


Figure A.1: Plot showing gaps in the estimation of number of pixels damaged by Cyclone Sidr caused by the Landsat 7 Scan Line Corrector failure. While the CCDC algorithm uses temporal modelling to interpolate between observations, the change detection process itself relies on real observations being available which can be compared to the fitted model.

Table A.1: Mangrove area lost and gained for each year in comparison to the previous year (km^2).

Year	Change	Net Loss	Net Gain	Total
1988	-	-	-	6368.3
1989	39.9	86.7	126.6	6408.2
1990	171.8	19.8	191.6	6580.0
1991	32.6	11.9	44.5	6612.6
1992	15.2	5.6	20.8	6627.8
1993	7.7	3.7	11.4	6635.5
1994	2.9	4.5	7.4	6638.4
1995	10.7	4.7	15.4	6649.1
1996	2.0	5.7	7.7	6651.1
1997	5.0	7.0	12.0	6656.1
1998	4.4	6.7	11.1	6660.5
1999	2.9	7.0	9.9	6663.4
2000	2.5	6.5	9.0	6665.9
2001	0.4	10.4	10.8	6666.3
2002	−0.7	12.3	11.6	6665.6

Table A.1: *Cont.*

Year	Change	Net Loss	Net Gain	Total
2003	1.5	12.9	14.4	6667.1
2004	5.5	11.9	17.4	6672.6
2005	−3.5	12.4	8.9	6669.1
2006	−3.0	10.6	7.6	6666.1
2007	−90.1	105.0	14.9	6576.0
2008	82.5	21.7	104.2	6658.5
2009	−2.5	17.4	14.9	6656.0
2010	9.2	9.6	18.8	6665.2
2011	−2.0	6.4	4.4	6663.2
2012	−3.2	7.0	3.8	6660.0
2013	−3.8	9.6	5.8	6656.2
2014	−7.1	14.1	7.0	6649.1
2015	−3.4	10.0	6.6	6645.7
2016	−40.6	43.8	3.2	6605.1
2017	12.5	10.7	23.2	6617.6

References

[1] Abdul Aziz, A., Phinn, S., and Dargusch, P., "Investigating the decline of ecosystem services in a production mangrove forest using Landsat and object-based image analysis," *Estuarine, Coastal and Shelf Science*, vol. 164, pp. 353–366, 2015. DOI: 10.1016/j.ecss.2015.07.047.

[2] Abdul Aziz, A., Phinn, S., Dargusch, P., Omar, H., and Arjasakusuma, S., "Assessing the potential applications of Landsat image archive in the ecological monitoring and management of a production mangrove forest in Malaysia," *Wetlands Ecology and Management*, vol. 23, no. 6, pp. 1049–1066, 2015. DOI: 10.1007/s11273-015-9443-1.

[3] Acharya, D., Rani, A., Agarwal, S., and Singh, V., "Application of adaptive Savitzky–Golay filter for EEG signal processing," *Perspectives in Science*, vol. 8, pp. 677–679, 2016. DOI: 10.1016/j.pisc.2016.06.056.

[4] Adole, T., Dash, J., and Atkinson, P. M., "Characterising the land surface phenology of Africa using 500 m MODIS EVI," *Applied Geography*, vol. 90, no. December 2017, pp. 187–199, 2018. DOI: 10.1016/j.apgeog.2017.12.006.

[5] Ahmed, O. S., Wulder, M. A., White, J. C., Hermosilla, T., Coops, N. C., and Franklin, S. E., "Classification of annual non-stand replacing boreal forest change in Canada using Landsat time series: a case study in northern Ontario," *Remote Sensing Letters*, vol. 8, no. 1, pp. 29–37, 2017. DOI: 10.1080/2150704X.2016.1233371.

[6] Akhter, M., Iqbal, Z., and Chowdhury, R. M., "ASTER imagery of forest areas of Sundarban damaged by cyclone Sidr," *ISME/GLOMIS Electronic Journal*, vol. 6, no. 1, 2008.

[7] AlMahamdy, M. and Riley, H. B., "Performance study of different denoising methods for ECG signals," in *Procedia Computer Science*, vol. 37, Elsevier, 2014, pp. 325–332. DOI: 10.1016/j.procs.2014.08.048.

[8] Almahasheer, H., Aljowair, A., Duarte, C. M., and Irigoien, X., "Decadal stability of Red Sea mangroves," *Estuarine, Coastal and Shelf Science*, vol. 169, pp. 164–172, 2016. DOI: 10.1016/J.ECSS.2015.11.027.

[9] Alongi, D. M., "Present state and future of the world's mangrove forests," *Environmental Conservation*, vol. 29, no. 3, pp. 331–349, 2002. DOI: 10.1017/S0376892902000231.

[10] Anthony, E. J. *et al.*, "The Amazon-influenced muddy coast of South America: A review of mud-bank-shoreline interactions," *Earth-Science Reviews*, vol. 103, no. 3-4, pp. 99–121, 2010. DOI: 10.1016/j.earscirev.2010.09.008.

[11] Anwar, M. S. and Takewaka, S., "Analyses on phenological and morphological variations of mangrove forests along the southwest coast of Bangladesh," *Journal of Coastal Conservation*, vol. 18, no. 4, pp. 339–357, 2014. DOI: 10.1007/s11852-014-0321-4.

[12] Arbeláez, P., Maire, M., Fowlkes, C., and Malik, J., "Contour Detection and Hierarchical Image Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 5, pp. 898–916, 2011. DOI: 10.1109/TPAMI.2010.161.

[13] Asbridge, E., Lucas, R., Ticehurst, C., and Bunting, P., "Mangrove response to environmental change in Australia's Gulf of Carpentaria," *Ecology and Evolution*, vol. 6, no. 11, pp. 3523–3539, 2016. DOI: 10.1002/ece3.2140.

[14] Aslan, A., Rahman, A. F., Robeson, S. M., and Ilman, M., "Land-use dynamics associated with mangrove deforestation for aquaculture and the subsequent abandonment of ponds," *Science of The Total Environment*, vol. 791, p. 148 320, 2021. DOI: https://doi.org/10.1016/j.scitotenv.2021.148320.

[15] Awty-Carroll, K. "COLD: Continuous Monitoring of Land Disturbance using ODC/NetCDF files," Aberystwyth University. (2020), [Online]. Available: https://github.com/klh5/COLD (visited on 04/19/2022).

[16] Awty-Carroll, K., *Simulated NDVI time series repository*, 2019. DOI: 10.17605/OSF.IO/TAF9Y.

[17] Awty-Carroll, K., Bunting, P., Hardy, A., and Bell, G., "Using Continuous Change Detection and Classification of Landsat Data to Investigate Long-Term Mangrove Dynamics in the Sundarbans Region," *Remote Sensing*, vol. 11, no. 23, p. 2833, 2019. DOI: 10.3390/rs11232833.

[18] Ayers, D., Oliver, S., Woodcock, R., and Contributors, *Open Data Cube*, 2019.

[19] Azzari, G and Lobell, D. B., "Landsat-based classification in the cloud: An opportunity for a paradigm shift in land cover monitoring," *Remote Sensing of Environment*, 2017. DOI: 10.1016/j.rse.2017.05.025.

[20] Badeck, F. W. *et al.*, *Responses of spring phenology to climate change*, 2004. DOI: 10.1111/j.1469-8137.2004.01059.x.

[21] Bai, J. and Perron, P., "Computation and analysis of multiple structural change models," *Journal of Applied Econometrics*, vol. 18, no. 1, pp. 1–22, 2003. DOI: 10.1002/jae.659.

[22] Baird, J., "Oil's shame in africa," *Newsweek (Atlantic Edition)*, vol. 156, no. 4, p. 16, 2010.

[23] Bandaranayake, W. M., "Traditional and medicinal uses of mangroves," *Mangroves and Salt Marshes*, vol. 2, no. 3, pp. 133–148, 1998. DOI: 10.1023/A:1009988607044.

[24] Banerjee, K., Gatti, R. C., and Mitra, A., "Climate change-induced salinity variation impacts on a stenoecious mangrove species in the Indian Sundarbans," *Ambio*, vol. 46, no. 4, pp. 492–499, 2017. DOI: 10.1007/s13280-016-0839-9.

[25] Beck, P. S., Jönsson, P., Høgda, K. A., Karlsen, S. R., Eklundh, L, and Skidmore, A. K., "A ground-validated NDVI dataset for monitoring vegetation dynamics and mapping phenology in Fennoscandia and the Kola peninsula," *International Journal of Remote Sensing*, vol. 28, no. 19, pp. 4311–4330, 2007. DOI: 10.1080/01431160701241936.

[26] Beck, P. S., Atzberger, C., Høgda, K. A., Johansen, B., and Skidmore, A. K., "Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI," *Remote Sensing of Environment*, vol. 100, no. 3, pp. 321–334, 2006. DOI: 10.1016/j.rse.2005.10.021. arXiv: /ehis.ebscohost.com/ [http:].

[27] Belward, A. S. and Skøien, J. O., "Who launched what, when and why; trends in global land-cover observation capacity from civilian earth observation satellites," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 103, pp. 115–128, 2015. DOI: 10.1016/j.isprsjprs.2014.03.009.

[28] Bhowmik, A. and Cabral, P., "Damage and post-cyclone regeneration assessment of the Sundarbans botanic biodiversity caused by the Cyclone Sidr," *Sustainability*, vol. 3, pp. 1–16, 2011.

[29] Bischof, H., Pinz, A. J., and Schneiden, W., "Multispectral Classification of Landsat-Images Using Neural Networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, no. 3, pp. 482–490, 1992. DOI: 10.1109/36.142926.

[30] Bonney, M. T., Danby, R. K., and Treitz, P. M., "Landscape variability of vegetation change across the forest to tundra transition of central Canada," *Remote Sensing of Environment*, vol. 217, pp. 18–29, 2018. DOI: 10.1016/J.RSE.2018.08.002.

[31] Boschetti, L., Roy, D. P., Justice, C. O., and Humber, M. L., "MODIS-Landsat fusion for large area 30m burned area mapping," *Remote Sensing of Environment*, vol. 161, pp. 27–42, 2015. DOI: 10.1016/j.rse.2015.01.022.

[32] Boschetti, M., Stroppiana, D., Brivio, P. A., and Bocchi, S., "Multi-year monitoring of rice crop phenology through time series analysis of MODIS images," *International Journal of Remote Sensing*, vol. 30, no. 18, pp. 4643–4662, 2009. DOI: 10.1080/01431160802632249.

[33] Boyd, D. S. and Danson, F. M., *Satellite remote sensing of forest resources: Three decades of research development*, 2005. DOI: 10.1191/0309133305pp432ra.

[34] Brooks, E. B., Thomas, V. A., Wynne, R. H., and Coulston, J. W., "Fitting the multitemporal curve: A fourier series approach to the missing data problem in remote sensing analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 9, pp. 3340–3353, 2012. DOI: 10.1109/TGRS.2012.2183137.

[35] Brooks, E. B., Wynne, R. H., Thomas, V. A., Blinn, C. E., and Coulston, J. W., "On-the-fly massively multitemporal change detection using statistical quality control charts and landsat data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 6, pp. 3316–3332, 2014. DOI: 10.1109/TGRS.2013.2272545.

[36] Brooks, E. B., Wynne, R. H., Thomas, V. A., Blinn, C. E., and Coulston, J., *Exponentially Weighted Moving Average Change Detection - Script and Sample Data*, 2014. DOI: https://doi.org/10.7294/W4WD3XHK.

[37] Brooks, E. B., Yang, Z., Thomas, V. A., and Wynne, R. H., "Edyn: Dynamic signaling of changes to forests using exponentially weighted moving average charts," *Forests*, vol. 8, no. 9, p. 304, 2017. DOI: 10.3390/f8090304.

[38] Bunting, P. and Clewley, D., *Atmospheric and Radiometric Correction of Satellite Imagery (ARCSI)*, 2018.

[39] Bunting, P. *et al.*, "The global mangrove watch - A new 2010 global baseline of mangrove extent," *Remote Sensing*, vol. 10, no. 10, p. 1669, 2018. DOI: 10.3390/rs10101669.

[40] Butt, B., Turner, M. D., Singh, A., and Brottem, L., "Use of MODIS NDVI to evaluate changing latitudinal gradients of rangeland phenology in Sudano-Sahelian West Africa," *Remote Sensing of Environment*, vol. 115, pp. 3367–3376, 2011. DOI: 10.1016/j.rse.2011.08.001.

[41] Cai, S. and Liu, D., "Detecting change dates from dense satellite time series using a sub-annual change detection algorithm," *Remote Sensing*, vol. 7, pp. 8705–8727, 2015. DOI: 10.3390/rs70708705.

[42] Çelik, T., "Bayesian change detection based on spatial sampling and Gaussian mixture model," *Pattern Recognition Letters*, vol. 32, no. 12, pp. 1635–1642, 2011. DOI: 10.1016/j.patrec.2011.05.008.

[43] Chamberlain, D., Phinn, S., and Possingham, H., "Remote Sensing of Mangroves and Estuarine Communities in Central Queensland, Australia," *Remote Sensing*, vol. 12, no. 1, p. 197, 2020. DOI: 10.3390/rs12010197.

[44] Che, X. *et al.*, "Mapping extent dynamics of small lakes using downscaling MODIS surface reflectance," *Remote Sensing*, vol. 9, pp. 82–103, 2017. DOI: 10.3390/rs9010082.

[45] Chen, L., Michishita, R., and Xu, B., "Abrupt spatiotemporal land and water changes and their potential drivers in Poyang Lake, 2000-2012," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 98, pp. 85–93, 2014. DOI: 10.1016/j.isprsjprs.2014.09.014.

[46] Chen, S. H. and Pollino, C. A., "Good practice in Bayesian network modelling," *Environmental Modelling and Software*, vol. 37, pp. 134–145, 2012. DOI: 10.1016/j.envsoft.2012.03.012.

[47] Chong, V. C., *Sustainable utilization and management of Mangrove ecosystems of Malaysia*, 2006. DOI: 10.1080/14634980600717084.

[48] Clive A, M. *et al.*, "Forest loss and Borneo's climate," *Environmental Research Letters*, vol. 13, 2018.

[49] Cohen, W. B., Yang, Z., and Kennedy, R., "Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync - Tools for calibration and validation," *Remote Sensing of Environment*, vol. 114, no. 12, pp. 2911–2924, 2010. DOI: 10.1016/j.rse.2010.07.010.

[50] Cohen, W. B. *et al.*, "How Similar Are Forest Disturbance Maps Derived from Different Landsat Time Series Algorithms?" *Forests*, vol. 8, no. 4, p. 98, 2017. DOI: 10.3390/f8040098.

[51] Colwell, R. N., "Remote sensing of natural resources," *Remote Sensing of Natural Resources*, vol. 218, no. 1, pp. 54–71, 1968. DOI: 10.1201/b15159.

[52] Commons, C. "Attribution-sharealike 3.0 unported." (), [Online]. Available: https://creativecommons.org/licenses/by-sa/3.0/deed.en (visited on 05/31/2021).

[53] Coppin, P. R. and Bauer, M. E., "Digital Change Detection in Forest Ecosystems with Remote Sensing Imagery," *Remote Sensing Reviews*, vol. 13, no. 3-4, pp. 207–234, 1996. DOI: 10.1080/02757259609532305.

[54] ——, "Processing of Multitemporal Landsat TM Imagery to Optimize Extraction of Forest Cover Change Features," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, no. 4, pp. 918–927, 1994. DOI: 10.1109/36.298020.

[55] Cornforth, W. A., Fatoyinbo, T. E., Freemantle, T. P., and Pettorelli, N., "Advanced land observing satellite phased array type L-Band SAR (ALOS PALSAR) to inform the conservation of mangroves: Sundarbans as a case study," *Remote Sensing*, vol. 5, no. 1, pp. 224–237, 2013. DOI: 10.3390/rs5010224.

[56] Cox, R. T., "Probability, Frequency and Reasonable Expectation," *American Journal of Physics*, vol. 14, no. 1, pp. 1–13, 1946. DOI: 10.1119/1.1990764.

[57] Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., and Bharath, A. A., "Generative adversarial network: An overview," *IEEE Signal Processing Magazine*, pp. 53–65, 2018. DOI: 10.19650/j.cnki.cjsi.J1804413.

[58] D'Addabbo, A., Refice, A., Pasquariello, G., Lovergine, F. P., Capolongo, D., and Manfreda, S., "A Bayesian Network for Flood Detection Combining SAR Imagery and Ancillary Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 6, pp. 3612–3625, 2016. DOI: 10.1109/TGRS.2016.2520487.

[59] Darmawan, Y., "Comparison of the Vegetation Indices To Detect the Tropical Rain Forest Changes Using BFAST Model," *International Journal of Remote Sensing and Earth Sciences (IJReSES)*, vol. 9, no. 1, pp. 21–34, 2012. DOI: 10.30536/j.ijreses.2012.v9.a1823.

[60] Datta, D. and Deb, S., "Analysis of coastal land use/land cover changes in the Indian Sunderbans using remotely sensed data," *Geo-Spatial Information Science*, vol. 15, no. 4, pp. 241–250, 2012. DOI: 10.1080/10095020.2012.714104. arXiv: arXiv:1011.1669v3.

[61] ——, "Forest structure and soil properties of mangrove ecosystems under different management scenarios: Experiences from the intensely humanized landscape of Indian Sunderbans," *Ocean and Coastal Management*, vol. 140, pp. 22–33, 2017. DOI: 10.1016/j.ocecoaman.2017.02.022.

[62] Davis, C. L., Hoffman, M. T., and Roberts, W., "Long-term trends in vegetation phenology and productivity over Namaqualand using the GIMMS AVHRR NDVI3g data from 1982 to 2011," *South African Journal of Botany*, vol. 111, pp. 76–85, 2017. DOI: 10.1016/j.sajb.2017.03.007.

[63] De Beer, T., Burggraeve, A., Fonteyne, M., Saerens, L., Remon, J. P., and Vervaet, C., "Near infrared and Raman spectroscopy for the in-process monitoring of pharmaceutical production processes," *International Journal of Pharmaceutics*, vol. 417, no. 1-2, pp. 32–47, 2011. DOI: 10.1016/j.ijpharm.2010.12.012.

[64] DeVries, B., Decuyper, M., Verbesselt, J., Zeileis, A., Herold, M., and Joseph, S., "Tracking disturbance-regrowth dynamics in tropical forests using structural change detection and Landsat time series," *Remote Sensing of Environment*, vol. 169, pp. 320–334, 2015. DOI: 10.1016/j.rse.2015.08.020.

[65] DeVries, B., Verbesselt, J., Kooistra, L., and Herold, M., "Robust monitoring of small-scale forest disturbances in a tropical montane forest using Landsat time series," *Remote Sensing of Environment*, vol. 161, no. March, pp. 107–121, 2015. DOI: 10.1016/j.rse.2015.02.012.

[66] Deng, C. and Zhu, Z., "Continuous subpixel mapping of impervious surface area using Landsat time series," *Remote Sensing of Environment*, 2018. DOI: S0034425718304590.

[67] Dice, L. R., "Measures of the Amount of Ecologic Association Between Species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.

[68] Donato, D. C., Kauffman, J. B., Murdiyarso, D., Kurnianto, S., Stidham, M., and Kanninen, M., "Mangroves among the most carbon-rich forests in the tropics," *Nature Geoscience*, vol. 4, no. 5, pp. 293–297, 2011. DOI: 10.1038/ngeo1123.

[69] Duke, N. C. *et al.*, "Large-scale dieback of mangroves in Australia's Gulf of Carpentaria: A severe ecosystem response, coincidental with an unusually extreme weather event," *Marine and Freshwater Research*, vol. 68, no. 10, pp. 1816–1829, 2017. DOI: 10.1071/MF16322.

[70] Dutrieux, L. P., Verbesselt, J., Kooistra, L., and Herold, M., "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*, vol. 107, pp. 112–125, 2015. DOI: 10.1016/j.isprsjprs.2015.03.015.

[71] Eklundh, L., Johansson, T., and Solberg, S., "Mapping insect defoliation in Scots pine with MODIS time-series data," *Remote Sensing of Environment*, vol. 113, pp. 1566–1573, 2009. DOI: 10.1016/j.rse.2009.03.008.

[72] Esri. "Esri 2020 land cover," Esri. (2020), [Online]. Available: https://www.arcgis.com/home/item.html?id=d6642f8a4f6d4685a24ae2dc0c73d4ac (visited on 08/15/2021).

[73] European Space Agency. "Sentinel-2: Facts and Figures." Accessed: 7th March 2018. (2017), [Online]. Available: https://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Sentinel-2/Facts_and_figures.

[74] Feller, I. C., Friess, D. A., Krauss, K. W., and Lewis, R. R., "The state of the world's mangroves in the 21st century under climate change," *Hydrobiologia*, vol. 803, no. 1, pp. 1–12, 2017. DOI: 10.1007/s10750-017-3331-z.

[75] Forkel, M., Carvalhais, N., Verbesselt, J., Mahecha, M. D., Neigh, C. S., and Reichstein, M., "Trend Change detection in NDVI time series: Effects of inter-annual variability and methodology," *Remote Sensing*, vol. 5, no. 5, pp. 2113–2144, 2013. DOI: 10.3390/rs5052113.

[76] Fragal, E. H., Silva, T. S. F., and Novo, E. M. L. d. M., "Reconstructing historical forest cover change in the Lower Amazon floodplains using the LandTrendr algorithm," *Acta Amazonica*, vol. 46, no. 1, pp. 13–24, 2016. DOI: 10.1590/1809-4392201500835. arXiv: 0052010.

[77] Frazier, R. J., Coops, N. C., and Wulder, M. A., "Boreal Shield forest disturbance and recovery trends using Landsat time series," *Remote Sensing of Environment*, vol. 170, pp. 317–327, 2015. DOI: 10.1016/j.rse.2015.09.015.

[78] Freitas, R. M. and Shimabukuro, Y. E., "Combining wavelets and linear spectral mixture model for MODIS satellite sensor timeseries analysis," *Journal of Computational Interdisciplinary Sciences*, vol. 1, no. 1, pp. 33–38, 2008. DOI: 10.6062/jcis.2008.01.01.0005.

[79] Fromard, F, Vega, C, and Proisy, C, "Half a century of dynamic coastal change affecting mangrove shorelines of French Guiana. A case study based on remote sensing data analyses and field surveys," in *Marine Geology*, vol. 208, 2004, pp. 265–280. DOI: 10.1016/j.margeo.2004.04.018.

[80] Fu, P. and Weng, Q., "A time series analysis of urbanization induced land use and land cover change and its impact on land surface temperature with Landsat imagery," *Remote Sensing of Environment*, vol. 175, pp. 205–214, 2016. DOI: 10.1016/j.rse.2015.12.040.

[81] Fu, Y., Chen, H., Niu, H., Zhang, S., and Yang, Y., "Spatial and temporal variation of vegetation phenology and its response to climate changes in Qaidam Basin from 2000 to 2015," *Journal of Geographical Sciences*, vol. 28, pp. 400–414, 2018. DOI: 10.1007/s11442-018-1480-2.

[82] Fujiyoshi, H., Hirakawa, T., and Yamashita, T., "Deep learning-based image recognition for autonomous driving," *IATSS Research*, vol. 43, no. 4, pp. 244–252, 2019. DOI: 10.1016/j.iatssr.2019.11.008.

[83] GDAL/OGR contributors. "GDAL/OGR geospatial data abstraction software library," Open Source Geospatial Foundation. (2019), [Online]. Available: https://gdal.org (visited on 07/30/2019).

[84] GERSL GitHub contributors. "Global Environmental Remote Sensing Laboratory repository." Accessed: 23rd September 2020, Global Environmental Remote Sensing Laboratory. (2020), [Online]. Available: https://github.com/GERSL.

[85] Galford, G. L., Mustard, J. F., Melillo, J., Gendrin, A., Cerri, C. C., and Cerri, C. E. P., "Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil," *Remote Sensing of Environment*, vol. 112, no. 2, pp. 576–587, 2008. DOI: 10.1016/j.rse.2007.05.017.

[86] Gao, Y., Ghilardi, A., Mas, J. F., Quevedo, A., Paneque-Gálvez, J., and Skutsch, M., "Assessing forest cover change in Mexico from annual MODIS VCF data (2000–2010)," *International Journal of Remote Sensing*, pp. 1–18, 2018. DOI: 10.1080/01431161.2018.1479789.

[87] Gardel, A. A. and Gratiot, N, "Monitoring of Coastal Dynamics in French Guiana from 16 Years of SPOT Satellite Images," *Journal of Coastal Research*, vol. 3, no. 39, pp. 1502–1505, 2006.

[88] Gaveau, D. L. *et al.*, "Four decades of forest persistence, clearance and logging on Borneo," *PLoS ONE*, vol. 9, no. 7, Bawa, K., Ed., e101654, 2014. DOI: 10.1371/journal.pone.0101654.

[89] Gensac, E, Lesourd, S, Gardel, A, Anthony, E. J., Proisy, C, and Loisel, H, "Short-term prediction of the evolution of mangrove surface areas: The example of the mud banks of Kourou and Sinnamary, French Guiana," *Journal of Coastal Research*, vol. 388, no. 64, pp. 388–392, 2011.

[90] Ghosh, A., Schmidt, S., Fickert, T., and Nüsser, M., "The Indian Sundarban mangrove forests: History, utilization, conservation strategies and local perception," *Diversity*, vol. 7, no. 2, pp. 149–169, 2015. DOI: 10.3390/d7020149.

[91] Ghosh, M. K., Kumar, L., and Roy, C., "Mapping long-term changes in mangrove species composition and distribution in the Sundarbans," *Forests*, vol. 7, no. 12, p. 305, 2016. DOI: 10.3390/f7120305.

[92] Ghosh, S. M. and Behera, M. D., "Aboveground biomass estimates of tropical mangrove forest using Sentinel-1 SAR coherence data - The superiority of deep learning over a semi-empirical model," *Computers and Geosciences*, vol. 150, no. October 2020, p. 104 737, 2021. DOI: 10.1016/j.cageo.2021.104737.

[93] Gillanders, S. N., Coops, N. C., Wulder, M. A., and Goodwin, N. R., "Application of Landsat satellite imagery to monitor land-cover changes at the Athabasca Oil Sands, Alberta, Canada," *Canadian Geographer*, vol. 52, no. 4, pp. 466–485, 2008. DOI: 10.1111/j.1541-0064.2008.00225.x.

[94] Giri, C. *et al.*, "Status and distribution of mangrove forests of the world using earth observation satellite data," *Global Ecology and Biogeography*, vol. 20, no. 1, pp. 154–159, 2011. DOI: 10.1111/j.1466-8238.2010.00584.x.

[95] Giri, C., *Observation and monitoring of mangrove forests using remote sensing: Opportunities and challenges*, 2016. DOI: 10.3390/rs8090783.

[96] Giri, C. and Muhlhausen, J., "Mangrove forest distributions and dynamics in Madagascar (1975-2005)," *Sensors*, vol. 8, pp. 2104–2117, 2008. DOI: 10.3390/s8042104.

[97] Giri, C., Pengra, B., Zhu, Z., Singh, A., and Tieszen, L. L., "Monitoring mangrove forest dynamics of the Sundarbans in Bangladesh and India using multitemporal satellite data from 1973 to 2000," *Estuarine, Coastal and Shelf Science*, vol. 73, no. 1-2, pp. 91–100, 2007. DOI: 10.1016/j.ecss.2006.12.019.

[98] Giri, C. *et al.*, "Distribution and dynamics of mangrove forests of South Asia," *Journal of Environmental Management*, vol. 148, pp. 101–111, 2015. DOI: 10.1016/j.jenvman.2014.01.020.

[99] Giri, S. *et al.*, "A study on abundance and distribution of mangrove species in Indian Sundarban using remote sensing technique," *Journal of Coastal Conservation*, vol. 18, no. 4, pp. 359–367, 2014. DOI: 10.1007/s11852-014-0322-3.

[100] Goessens, A. *et al.*, "Is Matang Mangrove Forest in Malaysia sustainably rejuvenating after more than a century of conservation and harvesting management?" *PLoS ONE*, vol. 9, no. 8, Heil, M., Ed., e105069, 2014. DOI: 10.1371/journal.pone.0105069.

[101] Goodwin, N. R. and Collett, L. J., "Development of an automated method for mapping fire history captured in Landsat TM and ETM+ time series across Queensland, Australia," *Remote Sensing of Environment*, vol. 148, pp. 206–221, 2014. DOI: 10.1016/j.rse.2014.03.021.

[102] Goodwin, N. R., Coops, N. C., Wulder, M. A., Gillanders, S., Schroeder, T. A., and Nelson, T., "Estimation of insect infestation dynamics using a temporal sequence of Landsat data," *Remote Sensing of Environment*, vol. 112, no. 9, pp. 3680–3689, 2008. DOI: 10.1016/j.rse.2008.05.005.

[103] Gopal, B. and Chauhan, M., "Biodiversity and its conservation in the Sundarban mangrove ecosystem," *Aquatic Sciences*, vol. 68, no. 3, pp. 338–354, 2006. DOI: 10.1007/s00027-006-0868-8.

[104] Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R., "Google Earth Engine: Planetary-scale geospatial analysis for everyone," *Remote Sensing of Environment*, vol. 202, pp. 18–27, 2017. DOI: 10.1016/j.rse.2017.06.031.

[105] Gouveia, C., Trigo, R. M., and DaCamara, C. C., "Drought and vegetation stress monitoring in Portugal using satellite data," *Natural Hazards and Earth System Science*, vol. 9, no. 1, pp. 185–195, 2009. DOI: 10.5194/nhess-9-185-2009.

[106] Green, E. P., Clark, C. D., Mumby, P. J., Edwards, A. J., and Ellis, A. C., "Remote sensing techniques for mangrove mapping," *International Journal of Remote Sensing*, vol. 19, no. 5, pp. 935–956, 1998. DOI: 10.1080/014311698215801.

[107] Griffiths, P., Kuemmerle, T., Kennedy, R. E., Abrudan, I. V., Knorn, J., and Hostert, P., "Using annual time-series of Landsat images to assess the effects of forest restitution in post-socialist Romania," *Remote Sensing of Environment*, vol. 118, pp. 199–214, 2012. DOI: 10.1016/j.rse.2011.11.006.

[108] Grogan, K., Pflugmacher, D., Hostert, P., Kennedy, R., and Fensholt, R., "Cross-border forest disturbance and the role of natural rubber in mainland Southeast Asia using annual Landsat time series," *Remote Sensing of Environment*, vol. 169, pp. 438–453, 2015. DOI: 10.1016/j.rse.2015.03.001.

[109] Grogan, K., Pflugmacher, D., Hostert, P., Verbesselt, J., and Fensholt, R., "Mapping clearances in tropical dry forests using breakpoints, trend, and seasonal components from modis time series: Does forest type matter?" *Remote Sensing*, vol. 8, pp. 657–684, 2016. DOI: 10.3390/rs8080657.

[110] Guay, K. C., Beck, P. S. A., Berner, L. T., Goetz, S. J., Baccini, A., and Buermann, W., "Vegetation productivity patterns at high northern latitudes: A multi-sensor satellite data assessment," *Global Change Biology*, vol. 20, no. 10, pp. 3147–3158, 2014. DOI: 10.1111/gcb.12647.

[111] Hamunyela, E., Verbesselt, J., and Herold, M., "Using spatial context to improve early detection of deforestation from Landsat time series," *Remote Sensing of Environment*, vol. 172, pp. 126–138, 2016. DOI: 10.1016/j.rse.2015.11.006.

[112] Harris, T *et al.*, "Climate drivers of the 2015 Gulf of Carpentaria mangrove dieback," Earth Systems, Climate Change Hub Technical Report No. 2, NESP Earth Systems, and Climate Change Hub, Australia., Tech. Rep., 2017.

[113] Healey, S. P. *et al.*, "Mapping forest change using stacked generalization: An ensemble approach," *Remote Sensing of Environment*, vol. 204, pp. 717–728, 2018. DOI: 10.1016/j.rse.2017.09.029.

[114] Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., and Hobart, G. W., "An integrated Landsat time series protocol for change detection and generation of annual gap-free surface reflectance composites," *Remote Sensing of Environment*, vol. 158, pp. 220–234, 2015. DOI: 10.1016/j.rse.2014.11.005.

[115] Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., Hobart, G. W., and Campbell, L. B., "Mass data processing of time series Landsat imagery: pixels to data products for forest monitoring," *International Journal of Digital Earth*, vol. 9, no. 11, pp. 1035–1054, 2016. DOI: 10.1080/17538947.2016.1187673.

[116] Heumann, B. W., "Satellite remote sensing of mangrove forests: Recent advances and future opportunities," *Progress in Physical Geography*, vol. 35, no. 1, pp. 87–108, 2011. DOI: 10.1177/0309133310385371.

[117] ——, "Satellite remote sensing of mangrove forests: Recent advances and future opportunities," *Progress in Physical Geography*, vol. 35, no. 1, pp. 87–108, 2011. DOI: 10.1177/0309133310385371.

[118] Hilker, T. *et al.*, "A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MODIS," *Remote Sensing of Environment*, vol. 113, no. 8, pp. 1613–1627, 2009. DOI: 10.1016/j.rse.2009.03.007.

[119] Holden, C. "Yet Another Time Series Model." Accessed: 23rd September 2020. (2017), [Online]. Available: https://yatsm.readthedocs.io/en/latest/.

[120] Hou, X., Gao, S., Niu, Z., and Xu, Z., "Extracting grassland vegetation phenology in North China based on cumulative SPOT-VEGETATION NDVI data," *International Journal of Remote Sensing*, vol. 35, no. 9, pp. 3316–3330, 2014. DOI: 10.1080/01431161.2014.903437.

[121] Houghton, R. A. and Nassikas, A. A., "Negative emissions from stopping deforestation and forest degradation, globally," *Global Change Biology*, vol. 24, no. 1, pp. 350–359, 2018. DOI: 10.1111/gcb.13876. arXiv: 0608246v3 [arXiv:physics].

[122] Huang, C., Goward, S. N., Schleeweis, K., Thomas, N., Masek, J. G., and Zhu, Z., "Dynamics of national forests assessed using the Landsat record: Case studies in eastern United States," *Remote Sensing of Environment*, vol. 113, no. 7, pp. 1430–1442, 2009. DOI: 10.1016/j.rse.2008.06.016.

[123] Huang, K., Zhou, T., and Zhao, X., "Extreme drought-induced trend changes in MODIS EVI time series in Yunnan, China," in *IOP Conference Series: Earth and Environmental Science*, vol. 17, IOP Publishing, 2014. DOI: 10.1088/1755-1315/17/1/012070.

[124] Hunter, J. S., "The Exponentially Weighted Moving Average," *Journal of Quality Technology*, vol. 18, no. 4, pp. 203–210, 1986. DOI: 10.1002/9780470400531.eorms0314.

[125] Hutchinson, J. M. S., Jacquin, A, Hutchinson, S. L., and Verbesselt, J, "Monitoring vegetation change and dynamics on U.S. Army training lands using satellite image time series analysis," *Journal of Environmental Management*, vol. 150, no. February, pp. 355–366, 2015. DOI: 10.1016/j.jenvman.2014.08.002.

[126] IPCC, "Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change," Tech. Rep., 2013. DOI: doi:10.1017/CBO9781107415324.

[127] Ibharim, N. A., Mustapha, M. A., Lihan, T., and Mazlan, A. G., "Mapping mangrove changes in the Matang Mangrove Forest using multi temporal satellite imageries," *Ocean and Coastal Management*, vol. 114, pp. 64–76, 2015. DOI: 10.1016/j.ocecoaman.2015.06.005.

[128] Ienco, D., Gaetano, R., Dupaquier, C., and Maurel, P., "Land Cover Classification via Multitemporal Spatial Data by Deep Recurrent Neural Networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 10, pp. 1685–1689, 2017. DOI: 10.1109/LGRS.2017.2728698. arXiv: 1704.04055.

[129] Iftekhar, M. S. and Islam, M., "Managing mangroves in Bangladesh: A strategy analysis," *Journal of Coastal Conservation*, vol. 10, no. 1, pp. 139–146, 2004. DOI: 10.1652/1400-0350(2004)010[0139:mmibas]2.0.co;2.

[130] Iftekhar, M. S. and Saenger, P., "Vegetation dynamics in the Bangladesh Sundarbans mangroves: A review of forest inventories," *Wetlands Ecology and Management*, vol. 16, no. 4, pp. 291–312, 2008. DOI: 10.1007/s11273-007-9063-5.

[131] Imhoff, M., Vermillion, C., Story, M., Khan, F., and Polcyn, F., "Forest Canopy Characterization and Vegetation Penetration Assessment with Space-Borne Radar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-24, no. 4, pp. 535–542, 1986. DOI: 10.1109/TGRS.1986.289668.

[132] Innocent Abbas, I., "An Assessment of Land Use/Land Cover Changes in a Section of Niger Delta, Nigeria," *Frontiers in Science*, vol. 2, no. 6, pp. 137–143, 2013. DOI: 10.5923/j.fs.20120206.02.

[133] Innocent Abbas, I. and Johnson Fasona, M., "Remote Sensing and Geographic Information Techniques: Veritable Tools for Land Degradation Assessment," *American Journal of Geographic Information System*, vol. 1, no. 1, pp. 1–6, 2012. DOI: 10.5923/j.ajgis.20120101.01.

[134] Islam, T. and Peterson, R. E., "Climatology of landfalling tropical cyclones in Bangladesh 1877-2003," *Natural Hazards*, vol. 48, no. 1, pp. 115–135, 2009. DOI: 10.1007/s11069-008-9252-4.

[135] Jamali, S., Seaquist, J., Eklundh, L., and Ardö, J., "Automated mapping of vegetation trends with polynomials using NDVI imagery over the Sahel," *Remote Sensing of Environment*, vol. 141, pp. 79–89, 2014. DOI: 10.1016/j.rse.2013.10.019.

[136] James, G. K., Adegoke, J. O., Osagie, S., Ekechukwu, S., Nwilo, P., and Akinyede, J., "Social valuation of mangroves in the Niger Delta region of Nigeria," *International Journal of Biodiversity Science, Ecosystem Services and Management*, vol. 9, no. 4, pp. 311–323, 2013. DOI: 10.1080/21513732.2013.842611.

[137] Jarron, L. R. *et al.*, "Differentiation of alternate harvesting practices using annual time series of landsat data," *Forests*, vol. 8, no. 1, p. 15, 2017. DOI: 10.3390/f8010015.

[138] Jeganathan, C., Dash, J., and Atkinson, P. M., "Remotely sensed trends in the phenology of northern high latitude terrestrial vegetation, controlling for land cover change and vegetation type," *Remote Sensing of Environment*, vol. 143, pp. 154–170, 2014. DOI: 10.1016/j.rse.2013.11.020.

[139] Jia, X. *et al.*, "Predict Land Covers with Transition Modeling and Incremental Learning," in *Proceedings of the 2017 SIAM International Conference on Data Mining*, Chawla, N. and Wang, W., Eds., Philadelphia, PA: Society for Industrial and Applied Mathematics, 2017, pp. 171–179. DOI: 10.1137/1.9781611974973.

[140] Jong, R. de, Verbesselt, J., Schaepman, M. E., and Bruin, S. de, "Trend changes in global greening and browning: Contribution of short-term trends to longer-term change," *Global Change Biology*, vol. 18, no. 2, pp. 642–655, 2012. DOI: 10.1111/j.1365-2486.2011.02578.x.

[141] Jönsson, A. M., Eklundh, L., Hellström, M., Bärring, L., and Jönsson, P., "Annual changes in MODIS vegetation indices of Swedish coniferous forests in relation to snow dynamics and tree phenology," *Remote Sensing of Environment*, vol. 114, no. 11, pp. 2719–2730, 2010. DOI: 10.1016/j.rse.2010.06.005.

[142] Jönsson, P. and Eklundh, L., "Seasonality extraction by function fitting to time-series of satellite sensor data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 8, pp. 1824–1832, 2002. DOI: 10.1109/TGRS.2002.802519. arXiv: TGRS.2002.802519 [10.1109].

[143] ——, "TIMESAT - A program for analyzing time-series of satellite sensor data," *Computers and Geosciences*, vol. 30, pp. 833–845, 2004. DOI: 10.1016/j.cageo.2004.05.006.

[144] Ju, J. and Roy, D. P., "The availability of cloud-free Landsat ETM+ data over the conterminous United States and globally," *Remote Sensing of Environment*, vol. 112, no. 3, pp. 1196–1211, 2008. DOI: 10.1016/j.rse.2007.08.011.

[145] Julien, Y. and Sobrino, J. A., "Global land surface phenology trends from GIMMS database," *International Journal of Remote Sensing*, vol. 30, no. 13, pp. 3495–3513, 2009. DOI: 10.1080/01431160802562255.

[146] Jönsson, P and Eklundh, L. "TIMESAT - A software package to analyse time-series of satellite sensor data." Accessed: 23rd September 2020. (2013-), [Online]. Available: http://web.nateko.lu.se/timesat/timesat.asp.

[147] ——, "TIMESAT 3.3 with seasonal trend decomposition and parallel processing - Software Manual." Accessed: 23rd September 2020. (2017), [Online]. Available: http://web.nateko.lu.se/timesat/docs/TIMESAT33_SoftwareManual.pdf.

[148] Kasper, D., Weidl, G., Dang, T., Breuel, G., Tamke, A., and Rosenstiel, W., "Object-oriented Bayesian networks for detection of lane change maneuvers," in *IEEE Intelligent Vehicles Symposium*, 2012, pp. 19–31. DOI: 10.1109/IVS.2011.5940468.

[149] Kennedy, R. E., Cohen, W. B., and Schroeder, T. A., "Trajectory-based change detection for automated characterization of forest disturbance dynamics," *Remote Sensing of Environment*, vol. 110, no. 3, pp. 370–386, 2007. DOI: 10.1016/j.rse.2007.03.010.

[150] Kennedy, R. E., Yang, Z., and Cohen, W. B., "Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr - Temporal segmentation algorithms," *Remote Sensing of Environment*, vol. 114, no. 12, pp. 2897–2910, 2010. DOI: 10.1016/j.rse.2010.07.008.

[151] Kennedy, R. E., Yang, Z., and Cohen, W. B. "LandTrendr-2012." Accessed: 23rd September 2020. (2017), [Online]. Available: https://github.com/KennedyResearch/LandTrendr-2012.

[152] Kennedy, R. E., Yang, Z., Cohen, W. B., Pfaff, E., Braaten, J., and Nelson, P., "Spatial and temporal patterns of forest disturbance and regrowth within the area of the Northwest Forest Plan," *Remote Sensing of Environment*, vol. 122, pp. 117–133, 2012. DOI: 10.1016/j.rse.2011.09.024.

[153] Kennedy, R. E. *et al.*, "Attribution of disturbance change agent from Landsat time-series in support of habitat monitoring in the Puget Sound region, USA," *Remote Sensing of Environment*, vol. 166, pp. 271–285, 2015. DOI: 10.1016/j.rse.2015.05.005.

[154] Kennedy, R. E. *et al.*, *Implementation of the LandTrendr algorithm on Google Earth Engine*, 2018. DOI: 10.3390/rs10050691.

[155] Khalil, G. M., "Cyclones and storm surges in Bangladesh: Some mitigative measures," *Natural Hazards*, vol. 6, no. 1, pp. 11–24, 1992. DOI: 10.1007/BF00162096.

[156] Krauss, K. W., From, A. S., Doyle, T. W., Doyle, T. J., and Barry, M. J., "Sea-level rise and landscape change influence mangrove encroachment onto marsh in the Ten Thousand Islands region of Florida, USA," *Journal of Coastal Conservation*, vol. 15, no. 4, pp. 629–638, 2011. DOI: 10.1007/s11852-011-0153-4.

[157] Krieter, J., Engler, J., Tölle, K. H., Timm, H. H., and Hohls, E., "Control charts applied to simulated sow herd datasets," *Livestock Science*, vol. 121, pp. 281–287, 2008. DOI: 10.1016/j.livsci.2008.06.025.

[158] Kristensen, E., Bouillon, S., Dittmar, T., and Marchand, C., "Organic carbon dynamics in mangrove ecosystems: A review," *Aquatic Botany*, vol. 89, no. 2, pp. 201–219, 2008. DOI: 10.1016/j.aquabot.2007.12.005.

[159] Kuenzer, C., Beijma, S. van, Gessner, U., and Dech, S., "Land surface dynamics and environmental challenges of the Niger Delta, Africa: Remote sensing-based analyses spanning three decades (1986-2013)," *Applied Geography*, vol. 53, pp. 354–368, 2014. DOI: 10.1016/j.apgeog.2014.07.002.

[160] Kuenzer, C., Bluemel, A., Gebhardt, S., Quoc, T. V., and Dech, S., "Remote sensing of mangrove ecosystems: A review," *Remote Sensing*, vol. 3, no. 5, pp. 878–928, 2011. DOI: 10.3390/rs3050878.

[161] Langner, A., Miettinen, J., and Siegert, F., "Land cover change 2002-2005 in Borneo and the role of fire derived from MODIS imagery," *Global Change Biology*, vol. 13, no. 11, pp. 2329–2340, 2007. DOI: 10.1111/j.1365-2486.2007.01442.x.

[162] Lau, K.-M and Weng, H., "Climate Signal Detection Using Wavelet Transform: How to Make a Time Series Sing," *Bulletin of the American Meteorological Society*, vol. 76, no. 12, pp. 2391–2402, 1995. DOI: 10.1175/1520-0477(1995)076<2391:CSDUWT>2.0.CO;2.

[163] Lee, G. *et al.* "PyWavelets - Wavelet Transforms in Python." Accessed: 23rd September 2020. (2006-), [Online]. Available: https://github.com/PyWavelets/pywt.

[164] Lehmann, E. A., Wallace, J. F., Caccetta, P. A., Furby, S. L., and Zdunic, K., "Forest cover trends from time series Landsat data for the Australian continent," *International Journal of Applied Earth Observation and Geoinformation*, vol. 21, pp. 453–462, 2012. DOI: 10.1016/j.jag.2012.06.005.

[165] Lewis, A. *et al.*, "Rapid, high-resolution detection of environmental change over continental scales from satellite data - the Earth Observation Data Cube," *International Journal of Digital Earth*, vol. 9, no. 1, pp. 106–111, 2016. DOI: 10.1080/17538947.2015.1111952.

[166] Lhermitte, S., Verbesselt, J., Verstraeten, W. W., Veraverbeke, S., and Coppin, P., "Assessing intra-annual vegetation regrowth after fire using the pixel based regeneration index," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 1, pp. 17–27, 2011. DOI: 10.1016/j.isprsjprs.2010.08.004.

[167] Lhermitte, S., Verbesselt, J., Verstraeten, W. W., and Coppin, P., "A Pixel Based Regeneration Index using Time Series Similarity and Spatial Context," *Photogrammetric Engineering and Remote Sensing*, vol. 76, no. 6, pp. 673–682, 2010. DOI: 10.14358/PERS.76.6.673.

[168] Li, H., Jia, M., Zhang, R., Ren, Y., and Wen, X., "Incorporating the plant phenological trajectory into mangrove species mapping with dense time series Sentinel-2 imagery and the Google Earth Engine platform," *Remote Sensing*, vol. 11, no. 21, 2019. DOI: 10.3390/rs11212479.

[169] Li, J., Deng, H., Li, P., and Yu, B., "Real-time infrared gas detection based on an adaptive Savitzky–Golay algorithm," *Applied Physics B: Lasers and Optics*, vol. 120, no. 2, pp. 207–216, 2015. DOI: 10.1007/s00340-015-6123-z.

[170] Li, J., Shen, Y., and Yang, C., "An adversarial generative network for crop classification from remote sensing timeseries images," *Remote Sensing*, vol. 13, no. 1, pp. 1–15, 2021. DOI: 10.3390/rs13010065.

[171] Liang, L., Hawbaker, T. J., Chen, Y., Zhu, Z., and Gong, P., "Characterizing recent and projecting future potential patterns of mountain pine beetle outbreaks in the Southern Rocky Mountains," *Applied Geography*, vol. 55, pp. 165–175, 2014. DOI: 10.1016/j.apgeog.2014.09.012.

[172] Liao, J., Zhen, J., Zhang, L., and Metternicht, G., "Understanding Dynamics of Mangrove Forest on Protected Areas of Hainan Island, China: 30 Years of Evidence from Remote Sensing," *Sustainability*, vol. 11, no. 19, p. 5356, 2019. DOI: 10.3390/su11195356.

[173] Lillesand, T., Kiefer, R. W., and Chipman, J., *Remote sensing and image interpretation*, 5th ed. John Wiley & Sons, Hoboken, New Jersey, 2004, ch. 7, pp. 494–507.

[174] Liu, S., Wei, X., Li, D., and Lu, D., "Examining forest disturbance and recovery in the subtropical forest region of Zhejiang Province using landsat time-series data," *Remote Sensing*, vol. 9, no. 5, p. 479, 2017. DOI: 10.3390/rs9050479.

[175] Liu, T., Yang, L., and Lunga, D., "Change detection using deep learning approach with object-based image analysis," *Remote Sensing of Environment*, vol. 256, no. January, p. 112 308, 2021. DOI: 10.1016/j.rse.2021.112308.

[176] Long, B. G. and Skewes, T. D., "A technique for mapping mangroves with Landsat TM satellite data and geographic information system," *Estuarine, Coastal and Shelf Science*, vol. 43, no. 3, pp. 373–381, 1996. DOI: 10.1006/ecss.1996.0076.

[177] Long, J., Giri, C., Primavera, J., and Trivedi, M., "Damage and recovery assessment of the Philippines' mangroves following Super Typhoon Haiyan," *Marine Pollution Bulletin*, vol. 109, no. 2, pp. 734–743, 2016. DOI: 10.1016/j.marpolbul.2016.06.080.

[178] Lovelock, C. E., Cahoon, D. R., Friess, D. A., Guntenspergen, G. R., Krauss, K. W., and Xuan, L., "The vulnerability of Indo-Pacific mangrove forests to sea-level rise," *Nature*, vol. 526, pp. 559–563, 2015. DOI: doi:10.1038/nature15538.

[179] Lucas, R. *et al.*, "Land cover mapping using digital earth Australia," *Data*, vol. 4, no. 4, 2019. DOI: 10.3390/data4040143.

[180] Lucas, R. *et al.*, "Structural characterisation of mangrove forests achieved through combining multiple sources of remote sensing data," *Remote Sensing of Environment*, vol. 237, p. 111 543, 2020. DOI: 10.1016/j.rse.2019.111543.

[181] Luo, C., Meng, S., Hu, X., Wang, X., and Zhong, Y., "Cropnet: Deep Spatial-Temporal-Spectral Feature Learning Network for Crop Classification from Time-Series Multi-Spectral Images," *International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 4187–4190, 2020. DOI: 10.1109/IGARSS39084.2020.9324097.

[182] Lymburner, L. *et al.*, "Mapping the multi-decadal mangrove dynamics of the Australian coastline," *Remote Sensing of Environment*, vol. 238, p. 111 185, 2020. DOI: 10.1016/j.rse.2019.05.004.

[183] Lyu, H., Lu, H., Mou, L., Lyu, H., Lu, H., and Mou, L., "Learning a Transferable Change Rule from a Recurrent Neural Network for Land Cover Change Detection," *Remote Sensing*, vol. 8, no. 6, p. 506, 2016. DOI: 10.3390/rs8060506.

[184] Malhotra, P., Vig, L., Shroff, G., and Agarwal, P., "Long Short Term Memory Networks for Anomaly Detection in Time Series," in *Proceedings of the 23rd European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, 2015, pp. 89–94.

[185] Mallat, S. G., "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989. DOI: 10.1109/34.192463. arXiv: NIHMS150003.

[186] Mandal, M. S. H. and Hosaka, T., "Assessing cyclone disturbances (1988–2016) in the Sundarbans mangrove forests using Landsat and Google Earth Engine," *Natural Hazards*, vol. 102, no. 1, pp. 133–150, 2020. DOI: 10.1007/s11069-020-03914-z.

[187] Markwardt, C. B. "MPFIT - Robust non-linear least squares curve fitting." Accessed: 23rd September 2020. (2017), [Online]. Available: http://cow.physics.wisc.edu/~craigm/idl/fitting.html.

[188] Marois, D. E. and Mitsch, W. J., "Coastal protection from tsunamis and cyclones provided by mangrove wetlands - A review," *International Journal of Biodiversity Science, Ecosystem Services and Management*, vol. 11, no. 1, pp. 71–83, 2015. DOI: 10.1080/21513732.2014.997292.

[189] Martínez, B. and Gilabert, M. A., "Vegetation dynamics from NDVI time series analysis using the wavelet transform," *Remote Sensing of Environment*, vol. 113, no. 9, pp. 1823–1842, 2009. DOI: 10.1016/j.rse.2009.04.016.

[190] Matgen, P., Hostache, R., Schumann, G., Pfister, L., Hoffmann, L., and Savenije, H. H. G., "Towards an automated SAR-based flood monitoring system: Lessons learned from two case studies," *Physics and Chemistry of the Earth*, vol. 36, pp. 241–252, 2011. DOI: 10.1016/j.pce.2010.12.009.

[191] Matosak, B. M., Maretto, R. V., Korting, T. S., Adami, M., and Fonseca, L. M., "Mapping Deforested Areas in the Cerrado Biome through Recurrent Neural Networks," *International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 1389–1392, 2020. DOI: 10.1109/IGARSS39084.2020.9324019.

[192] McCulloch, W. S. and Pitts, W., "A logical calculus of the ideas immanent in nervous activity," *The Bulletin of Mathematical Biophysics*, vol. 5, no. 4, pp. 115–133, 1943. DOI: 10.1007/BF02478259. arXiv: arXiv:1011.1669v3.

[193] McFeeters, S. K., "The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features," *International Journal of Remote Sensing*, vol. 17, no. 7, pp. 1425–1432, 1996. DOI: 10.1080/01431169608948714.

[194] McFeeters, S. K., "Using the normalized difference water index (ndwi) within a geographic information system to detect swimming pools for mosquito abatement: A practical approach," *Remote Sensing*, vol. 5, no. 7, pp. 3544–3561, 2013. DOI: 10.3390/rs5073544.

[195] Meigs, G. W., Kennedy, R. E., and Cohen, W. B., "A Landsat time series approach to characterize bark beetle and defoliator impacts on tree mortality and surface fuels in conifer forests," *Remote Sensing of Environment*, vol. 115, no. 12, pp. 3707–3718, 2011. DOI: 10.1016/j.rse.2011.09.009.

[196] Melendez-Pastor, I., Navarro-Pedreño, J., Koch, M., Gómez, I., and Hernández, E. I., "Land-cover phenologies and their relation to climatic variables in an anthropogenically impacted mediterranean coastal area," *Remote Sensing*, vol. 2, pp. 697–716, 2010. DOI: 10.3390/rs2030697.

[197] Mertens, K., Decuypere, E., De Baerdemaeker, J., and De Ketelaere, B., "Statistical control charts as a support tool for the management of livestock production," *Journal of Agricultural Science*, vol. 149, pp. 369–384, 2011. DOI: 10.1017/S0021859610001164.

[198] Meyer, M. C., "Inference using shape-restricted regression splines," *Annals of Applied Statistics*, vol. 2, no. 3, pp. 1013–1033, 2008. DOI: 10.1214/08-AOAS167. arXiv: 0811.1705.

[199] Meyer, M. C., Liao, X., Freeman, E., and Moisen, G. G. "ShapeSelectForest: Shape Selection for Landsat Time Series of Forest Dynamics." Accessed: 23rd September 2020. (2016), [Online]. Available: https://cran.r-project.org/web/packages/ShapeSelectForest/index.html.

[200] Mitchell, A. L., Rosenqvist, A., and Mora, B., *Current remote sensing approaches to monitoring forest degradation in support of countries measurement, reporting and verification (MRV) systems for REDD+*, 2017. DOI: 10.1186/s13021-017-0078-9.

[201] Mmom, P. C. and Arokoyu, S. B., "Mangrove forest depletion, biodiversity loss and traditional resources management practices in the Niger Delta, Nigeria," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 2, no. 1, pp. 28–34, 2010.

[202] Moisen, G. G. *et al.*, "Shape selection in Landsat time series: a tool for monitoring forest dynamics," *Global change biology*, vol. 22, no. 10, pp. 3518–3528, 2016. DOI: 10.1111/gcb.13358.

[203] Mondal, P., Liu, X., Fatoyinbo, T. E., and Lagomasino, D., "Evaluating combinations of sentinel-2 data and machine-learning algorithms for mangrove mapping in West Africa," *Remote Sensing*, vol. 11, no. 24, 2019. DOI: 10.3390/rs11242928.

[204] Morton, A. P. *et al.*, "The application of statistical process control charts to the detection and monitoring of hospital-acquired infections," *J. Qual. Clin. Practice (2001)*, vol. 21, no. 4, pp. 112–117, 2001. DOI: 10.1046/j.1440-1762.2001.00423.x.

[205] Moulin, S., Kergoat, L., Viovy, N., and Dedieu, G., "Global-scale assessment of vegetation phenology using NOAA/AVHRR satellite measurements," *Journal of Climate*, vol. 10, no. 6, pp. 1154–1170, 1997. DOI: 10.1175/1520-0442(1997)010<1154:GSAOVP>2.0.CO;2.

[206] Murillo-Sandoval, P. J., Hilker, T., Krawchuk, M. A., and Van Den Hoek, J., "Detecting and attributing drivers of forest disturbance in the Colombian andes using landsat time-series," *Forests*, vol. 9, 2018. DOI: 10.3390/f9050269.

[207] Nascimento, W. R., Souza-Filho, P. W. M., Proisy, C., Lucas, R. M., and Rosenqvist, A., "Mapping changes in the largest continuous Amazonian mangrove belt using object-based classification of multisensor satellite imagery," *Estuarine, Coastal and Shelf Science*, vol. 117, pp. 83–93, 2013. DOI: 10.1016/j.ecss.2012.10.005.

[208] Nature, I. U. f. C. of, *World Database on Protected Areas: Sundarbans*, 2019.

[209] Neogi, S. B. *et al.*, "Sundarban mangroves: diversity, ecosystem services and climate change impacts," *Asian Journal of Medical and Biological Research*, vol. 2, no. 4, pp. 488–507, 2017. DOI: 10.3329/ajmbr.v2i4.30988.

[210] Neuburger, J., Walker, K., Sherlaw-Johnson, C., Meulen, J. van der, and Cromwell, D. A., "Comparison of control charts for monitoring clinical performance using binary data," *BMJ Quality and Safety*, vol. 26, no. 11, pp. 919–928, 2017. DOI: 10.1136/bmjqs-2016-005526.

[211] Nguyen, T. H., Jones, S. D., Soto-Berelov, M., Haywood, A., and Hislop, S., "A spatial and temporal analysis of forest dynamics using Landsat time-series," *Remote Sensing of Environment*, vol. 217, pp. 461–475, 2018. DOI: 10.1016/j.rse.2018.08.028.

[212] O. Adegoke, J., Fageja, M., James, G., Agbaje, G., and E. Ologunorisa, T., "An Assessment of Recent Changes in the Niger Delta Coastline Using Satellite Imagery," *Journal of Sustainable Development*, vol. 3, no. 4, 2010. DOI: 10.5539/jsd.v3n4p277.

[213] Omar, H., Misman, M. A., and Linggok, V., "Characterizing and monitoring of mangroves in Malaysia using Landsat-based spatial-spectral variability," in *IOP Conference Series: Earth and Environmental Science*, vol. 169, IOP Publishing, 2018, p. 012 037. DOI: 10.1088/1755-1315/169/1/012037.

[214] Omo-Irabor, O. O., Olobaniyi, S. B., Akunna, J., Venus, V., Maina, J. M., and Paradzayi, C., "Mangrove vulnerability modelling in parts of Western Niger Delta, Nigeria using satellite images, GIS techniques and Spatial Multi-Criteria Analysis (SMCA)," *Environmental Monitoring and Assessment*, vol. 178, no. 1-4, pp. 39–51, 2011. DOI: 10.1007/s10661-010-1669-z.

[215] Onyena, A. P. and Sam, K., "A review of the threat of oil exploitation to mangrove ecosystem: Insights from Niger Delta, Nigeria," *Global Ecology and Conservation*, vol. 22, e00961, 2020. DOI: 10.1016/j.gecco.2020.e00961.

[216] Otero, V., Lucas, R., Van De Kerchove, R., Satyanarayana, B., Mohd-Lokman, H., and Dahdouh-Guebas, F., "Spatial analysis of early mangrove regeneration in the Matang Mangrove Forest Reserve, Peninsular Malaysia, using geomatics," *Forest Ecology and Management*, vol. 472, no. May, p. 118 213, 2020. DOI: 10.1016/j.foreco.2020.118213.

[217] Otero, V., Van De Kerchove, R., Satyanarayana, B., Mohd-Lokman, H., Lucas, R., and Dahdouh-Guebas, F., "An analysis of the early regeneration of mangrove forests using Landsat time series in the matang mangrove forest reserve, Peninsular Malaysia," *Remote Sensing*, vol. 11, no. 7, p. 774, 2019. DOI: 10.3390/rs11070774.

[218] Otero, V. *et al.*, "Managing mangrove forests from the sky: Forest inventory using field data and Unmanned Aerial Vehicle (UAV) imagery in the Matang Mangrove Forest Reserve, peninsular Malaysia," *Forest Ecology and Management*, vol. 411, no. December 2017, pp. 35–45, 2018. DOI: 10.1016/j.foreco.2017.12.049.

[219] Otter, D. W., Medina, J. R., and Kalita, J. K., "A Survey of the Usages of Deep Learning for Natural Language Processing," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 2, pp. 604–624, 2021. DOI: 10.1016/0004-3702(82)90032-7.

[220] Page, E. S., "Controlling the Standard Deviation by Cusums and Warning Lines," *Technometrics*, vol. 5, no. 3, pp. 307–315, 1963. DOI: 10.1080/00401706.1963.10490100.

[221] Palhares, R., D'Angelo, M., Loschi, R., and Takahashi, R., "Fuzzy/Bayesian change point detection approach to incipient fault detection," *IET Control Theory and Applications*, vol. 5, no. 4, pp. 539–551, 2011. DOI: 10.1049/iet-cta.2009.0033.

[222] Papadomanolaki, M., Vakalopoulou, M., and Karantzalos, K., "A Deep Multitask Learning Framework Coupling Semantic Segmentation and Fully Convolutional LSTM Networks for Urban Change Detection," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–18, 2021. DOI: 10.1109/TGRS.2021.3055584.

[223] Pasquarella, V. J., Bradley, B. A., and Woodcock, C. E., "Near-real-time monitoring of insect defoliation using Landsat time series," *Forests*, vol. 8, no. 8, p. 275, 2017. DOI: 10.3390/f8080275.

[224] Pasquarella, V. J., Holden, C. E., and Woodcock, C. E., "Improved mapping of forest type using spectral-temporal Landsat features," *Remote Sensing of Environment*, vol. 210, pp. 193–207, 2018. DOI: 10.1016/j.rse.2018.02.064.

[225] Pastor-Guzman, J., Dash, J., and Atkinson, P. M., "Remote sensing of mangrove forest phenology and its environmental drivers," *Remote Sensing of Environment*, vol. 205, pp. 71–84, 2018. DOI: 10.1016/j.rse.2017.11.009.

[226] Paul, B. K., "Why relatively fewer people died? The case of Bangladesh's cyclone sidr," *Natural Hazards*, vol. 50, no. 2, pp. 289–304, 2009. DOI: 10.1007/s11069-008-9340-5. arXiv: ISBN0921-030X,1573-0840.

[227] Pedregosa, F. *et al.*, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[228] Pesaran, M. H. and Timmermann, A., "Market timing and return prediction under model instability," *Journal of Empirical Finance*, vol. 9, no. 5, pp. 495–510, 2002. DOI: 10.1016/S0927-5398(02)00007-5.

[229] Pflugmacher, D., Cohen, W. B., Kennedy, R. E., and Yang, Z., "Using Landsat-derived disturbance and recovery history and lidar to map forest biomass dynamics," *Remote Sensing of Environment*, vol. 151, pp. 124–137, 2014. DOI: 10.1016/j.rse.2013.05.033.

[230] Pham, M. T., Courtrai, L., Friguet, C., Lefèvre, S., and Baussard, A., "YOLO-fine: One-stage detector of small objects under various backgrounds in remote sensing images," *Remote Sensing*, vol. 12, no. 15, pp. 1–26, 2020. DOI: 10.3390/RS12152501.

[231] Phompila, C., Lewis, M., Clarke, K., and Ostendorf, B., "Monitoring temporal Vegetation changes in Lao tropical forests," in *7th IGRSM International Remote Sensing and GIS Conference and Exhibition*, vol. 20, IOP Publishing, 2014. DOI: 10.1088/1755-1315/20/1/012054.

[232] Pickell, P. D., Hermosilla, T., Coops, N. C., Masek, J. G., Franks, S., and Huang, C., "Monitoring anthropogenic disturbance trends in an industrialized boreal forest with Landsat time series," *Remote Sensing Letters*, vol. 5, no. 9, pp. 783–792, 2014. DOI: 10.1080/2150704X.2014.967881.

[233] Pierce, J., "Mangroves special issue," *Journal of the Commonwealth Human Ecology Council*, vol. 30, pp. 44–49, Jun. 2020.

[234] Planet Labs Inc. "Planet homepage," Planet Labs Inc. (2021), [Online]. Available: https://www.planet.com/ (visited on 08/14/2021).

[235] Platt, R. V., Manthos, D., and Amos, J., "Estimating the Creation and Removal Date of Fracking Ponds Using Trend Analysis of Landsat Imagery," *Environmental Management*, vol. 61, no. 2, pp. 310–320, 2018. DOI: 10.1007/s00267-017-0983-4.

[236] Polidoro, B. A. *et al.*, "The loss of species: Mangrove extinction risk and geographic areas of global concern," *PLoS ONE*, vol. 5, no. 4, Hansen, D. M., Ed., pp. 1–10, 2010. DOI: 10.1371/journal.pone.0010095.

[237] Pontius Jr, R. G. and Millones, M., "Death to Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment," *International Journal of Remote Sensing*, vol. 32, no. 15, pp. 4407–4429, 2011. DOI: 10.1080/01431161.2011.552923.

[238] Potapov, P., Turubanova, S., and Hansen, M. C., "Regional-scale boreal forest cover and change mapping using Landsat data composites for European Russia," *Remote Sensing of Environment*, vol. 115, no. 2, pp. 548–561, 2011. DOI: 10.1016/j.rse.2010.10.001.

[239] Pouliot, D. and Latifovic, R., "Reconstruction of Landsat time series in the presence of irregular and sparse observations: Development and assessment in north-eastern Alberta, Canada," *Remote Sensing of Environment*, vol. 204, pp. 979–996, 2017. DOI: 10.1016/j.rse.2017.07.036.

[240] Proisy, C., Mougin, E., and Fromard, F., "Investigating correlations between radar data and mangrove forests characteristics," *International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 1, pp. 733–735, 1996. DOI: 10.1109/igarss.1996.516458.

[241] Proisy, C., Gratiot, N., Anthony, E. J., Gardel, A., Fromard, F., and Heuret, P., "Mud bank colonization by opportunistic mangroves: A case study from French Guiana using lidar data," *Continental Shelf Research*, vol. 29, no. 3, pp. 632–641, 2009. DOI: 10.1016/j.csr.2008.09.017.

[242] Rachmawani, D., Yulianda, F., Kusmana, C., and Boer, M., "Study of Mangroves Ecosystem Management at Binalatung in Tarakan City of North Kalimantan," *International Journal of Sciences: Basic and Applied Research (IJSBAR)*, vol. 26, no. 3, pp. 221–234, 2016.

[243] Raha, A., Das, S., Banerjee, K., and Mitra, A., "Climate change impacts on Indian Sunderbans: A time series analysis (1924-2008)," *Biodiversity and Conservation*, vol. 21, no. 5, pp. 1289–1307, 2012. DOI: 10.1007/s10531-012-0260-z.

[244] Rahman, A. F., Dragoni, D., and El-Masri, B., "Response of the Sundarbans coastline to sea level rise and decreased sediment flow: A remote sensing assessment," *Remote Sensing of Environment*, vol. 115, no. 12, pp. 3121–3128, 2011. DOI: 10.1016/j.rse.2011.06.019.

[245] Rapids AI. "Rapids: Open gpu data science," Rapids AI. (2021), [Online]. Available: https://rapids.ai/ (visited on 08/16/2021).

[246] Rees, W. G., Williams, M, and Vitebsky, P, "Mapping land cover change in a reindeer herding area of the Russian arctic using Landsat TM and ETM+ imagery and indigenous knowledge," *Remote Sensing of Environment*, vol. 85, no. 4, pp. 441–452, 2003. DOI: 10.1016/S0034-4257(03)00037-3.

[247] Refice, A., D'Addabbo, A., Pasquariello, G., Lovergine, F. P., Capolongo, D., and Manfreda, S., "Towards high-precision flood mapping: Multi-temporal SAR/InSAR data, Bayesian inference, and hydrologic modeling," *International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 2015-Novem, pp. 1381–1384, 2015. DOI: 10.1109/IGARSS.2015.7326034.

[248] Reiche, J., Bruin, S. de, Hoekman, D., Verbesselt, J., and Herold, M., "A Bayesian approach to combine landsat and ALOS PALSAR time series for near real-time deforestation detection," *Remote Sensing*, vol. 7, no. 5, pp. 4973–4996, 2015. DOI: 10.3390/rs70504973.

[249] Reiche, J., Hamunyela, E., Verbesselt, J., Hoekman, D., and Herold, M., "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*, vol. 204, pp. 147–161, 2018. DOI: 10.1016/j.rse.2017.10.034.

[250] Reiche, J., Verbesselt, J., Hoekman, D., and Herold, M., "Fusing Landsat and SAR time series to detect deforestation in the tropics," *Remote Sensing of Environment*, vol. 156, pp. 276–293, 2015. DOI: 10.1016/j.rse.2014.10.001.

[251] Revilla-Romero, B. *et al.*, "On the use of global flood forecasts and satellite-derived inundation maps for flood monitoring in data-sparse regions," *Remote Sensing*, vol. 7, no. 11, pp. 15 702–15 728, 2015. DOI: 10.3390/rs71115702.

[252] Rhif, M., Ben Abbes, A., Martinez, B., and Farah, I. R., "A deep learning approach for forecasting non-stationary big remote sensing time series," *Arabian Journal of Geosciences*, vol. 13, no. 22, 2020. DOI: 10.1007/s12517-020-06140-w.

[253] Richards, D. R. and Friess, D. A., "Rates and drivers of mangrove deforestation in Southeast Asia, 2000-2012," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 113, no. 2, pp. 344–349, 2016. DOI: 10.1073/pnas.1510272113.

[254] Richards, P. D. and VanWey, L., "Farm-scale distribution of deforestation and remaining forest cover in Mato Grosso," *Nature Climate Change*, vol. 6, no. 4, pp. 418–425, 2016. DOI: 10.1038/nclimate2854.

[255] Roerink, G. J., Menenti, M., and Verhoef, W., "Reconstructing cloudfree NDVI composites using Fourier analysis of time series," *International Journal of Remote Sensing*, vol. 21, no. 9, pp. 1911–1917, 2000. DOI: 10.1080/014311600209814.

[256] Rogers, K., Lymburner, L., Salum, R., Brooke, B. P., and Woodroffe, C. D., "Mapping of mangrove extent and zonation using high and low tide composites of Landsat data," *Hydrobiologia*, vol. 803, no. 1, pp. 49–68, 2017. DOI: 10.1007/s10750-017-3257-5.

[257] Roy, D. P. *et al.*, "Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity," *Remote Sensing of Environment*, vol. 185, pp. 57–70, 2016. DOI: 10.1016/j.rse.2015.12.024.

[258] Roy, D. P. *et al.*, "Landsat-8: Science and product vision for terrestrial global change research," *Remote Sensing of Environment*, vol. 145, pp. 154–172, 2014. DOI: 10.1016/j.rse.2014.02.001.

[259] Ruggieri, E., "A Bayesian approach to detecting change points in climatic records," *International Journal of Climatology*, vol. 33, no. 2, pp. 520–528, 2013. DOI: 10.1002/joc.3447.

[260] Saintilan, N., Wilson, N. C., Rogers, K., Rajkaran, A., and Krauss, K. W., "Mangrove expansion and salt marsh decline at mangrove poleward limits," *Global Change Biology*, vol. 20, no. 1, pp. 147–157, 2014. DOI: 10.1111/gcb.12341.

[261] Sakamoto, T., Yokozawa, M., Toritani, H., Shibayama, M., Ishitsuka, N., and Ohno, H., "A crop phenology detection method using time-series MODIS data," *Remote Sensing of Environment*, vol. 96, no. 3-4, pp. 366–374, 2005. DOI: 10.1016/j.rse.2005.03.008.

[262] Savitzky, A. and Golay, M. E. J., "Smoothing and Differentiation of Data by Simplified Least Squares Regression," *Analytical Chemistry*, vol. 36, no. 8, pp. 1627–1638, 1964.

[263] Saxena, R. *et al.*, "Towards a polyalgorithm for land use change detection," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 144, pp. 217–234, 2018. DOI: 10.1016/j.isprsjprs.2018.07.002.

[264] Schmidt, M., Lucas, R., Bunting, P., Verbesselt, J., and Armston, J., "Multi-resolution time series imagery for forest disturbance and regrowth monitoring in Queensland, Australia," *Remote Sensing of Environment*, vol. 158, no. March, pp. 156–168, 2015. DOI: 10.1016/j.rse.2014.11.015.

[265] Schneibel, A. *et al.*, "Assessment of spatio-temporal changes of smallholder cultivation patterns in the Angolan Miombo belt using segmentation of Landsat time series," *Remote Sensing of Environment*, vol. 195, pp. 118–129, 2017. DOI: 10.1016/j.rse.2017.04.012.

[266] Schroeder, T. A. *et al.*, "Testing a Landsat-based approach for mapping disturbance causality in U.S. forests," *Remote Sensing of Environment*, vol. 195, pp. 230–243, 2017. DOI: 10.1016/j.rse.2017.03.033.

[267] Schultz, M. *et al.*, "Forest Cover and Vegetation Degradation Detection in the Kavango Zambezi Transfrontier Conservation Area Using BFAST Monitor," *Remote Sensing*, vol. 10, no. 11, p. 1850, 2018. DOI: 10.3390/rs10111850.

[268] Seiler, R., Kogan, F, and Sullivan, J, "AVHRR-based vegetation and temperature condition indices for drought detection in Argentina," *Advances in Space Research*, vol. 21, no. 3, pp. 481–484, 1998. DOI: 10.1016/S0273-1177(97)00884-3.

[269] Sen, P. K., "Estimates of the Regression Coefficient Based on Kendall's Tau," *Journal of the American Statistical Association*, vol. 63, no. 324, pp. 1379–1389, 1968. DOI: 10.1080/01621459.1968.10480934.

[270] Senf, C., Pflugmacher, D., Wulder, M. A., and Hostert, P., "Characterizing spectral–temporal patterns of defoliator and bark beetle disturbances using Landsat time series," *Remote Sensing of Environment*, vol. 170, pp. 166–177, 2015. DOI: 10.1016/J.RSE.2015.09.019.

[271] Shao, Z., Tang, P., Wang, Z., Saleem, N., Yam, S., and Sommai, C., "BRRNet: A fully convolutional neural network for automatic building extraction from high-resolution remote sensing images," *Remote Sensing*, vol. 12, no. 6, pp. 1–17, 2020. DOI: 10.3390/rs12061050.

[272] Shaunak Ghosh. "Cracks in the shield: How the Sundarbans is dying and making Bengal prone to cyclones," Newslaundry. (2020), [Online]. Available: https://www.newslaundry.com/2020/12/18/cracks-in-the-shield-how-the-sundarbans-is-dying-and-making-bengal-prone-to-cyclones (visited on 04/16/2022).

[273] Shewhart, W. A., *Economic control of quality of manufactured product*. Lancaster Press, Lancaster, Pennsylvania, 1931.

[274] Shi, X., Chen, Z., Wang, H., Yeung, D. Y., Wong, W. K., and Woo, W. C., "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," *arXiv*, 2015. arXiv: arXiv:1506.04214v1.

[275] Sieber, A. *et al.*, "Landsat-based mapping of post-Soviet land-use change to assess the effectiveness of the Oksky and Mordovsky protected areas in European Russia," *Remote Sensing of Environment*, vol. 133, pp. 38–51, 2013. DOI: 10.1016/j.rse.2013.01.021.

[276] Silva, G. S., Schwartz, M., Morrison, R. B., and Linhares, D. C. L., "Monitoring breeding herd production data to detect PRRSV outbreaks," *Preventive Veterinary Medicine*, vol. 148, pp. 89–93, 2017. DOI: 10.1016/j.prevetmed.2017.10.012.

[277] Song, C., Huang, B., and You, S., "Comparison of three time-series NDVI reconstruction methods based on TIMESAT," *International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 2225–2228, 2012. DOI: 10.1109/IGARSS.2012.6351057.

[278] Souza, C. M. *et al.*, "Ten-year landsat classification of deforestation and forest degradation in the brazilian amazon," *Remote Sensing*, vol. 5, no. 11, pp. 5493–5513, 2013. DOI: 10.3390/rs5115493.

[279] Sra, S., Nowozin, S., and Wright, S., *Optimization for Machine Learning*, ser. Neural information processing series. MIT Press, 2012.

[280] Steiner, S. H. and Jones, M., "Risk-adjusted survival time monitoring with an updating exponentially weighted moving average (EWMA) control chart," *Statistics in Medicine*, vol. 29, no. 4, pp. 444–454, 2010. DOI: 10.1002/sim.3788.

[281] Suzuki, K., "Overview of deep learning in medical imaging," *Radiological Physics and Technology*, vol. 10, no. 3, pp. 257–273, 2017. DOI: 10.1007/s12194-017-0406-5.

[282] Symeonakis, E., Caccetta, P., Koukoulas, S., Furby, S., and Karathanasis, N., "Multi-temporal land-cover classification and change analysis with conditional probability networks: The case of Lesvos Island (Greece)," *International Journal of Remote Sensing*, vol. 33, no. 13, pp. 4075–4093, 2012. DOI: 10.1080/01431161.2011.640961.

[283] Tang, X., Bullock, E. L., Olofsson, P., Estel, S., and Woodcock, C. E., "Near real-time monitoring of tropical forest disturbance: New algorithms and assessment framework," *Remote Sensing of Environment*, vol. 224, pp. 202–218, 2019. DOI: 10.1016/j.rse.2019.02.003.

[284] Taureau, F., Robin, M., Proisy, C., Fromard, F., Imbert, D., and Debaine, F., "Mapping the Mangrove Forest Canopy Using Spectral Unmixing of Very High Spatial Resolution Satellite Images," *Remote Sensing*, vol. 11, no. 3, p. 367, 2019. DOI: 10.3390/rs11030367.

[285] Thomas, N., Bunting, P., Lucas, R., Hardy, A., Rosenqvist, A., and Fatoyinbo, T., "Mapping Mangrove Extent and Change: A Globally Applicable Approach," English, *Remote Sensing*, vol. 10, no. 9, p. 1466, Sep. 2018. DOI: 10.3390/rs10091466.

[286] Thomas, N., Lucas, R., Bunting, P., Hardy, A., Rosenqvist, A., and Simard, M., "Distribution and drivers of global mangrove forest change, 1996-2010," *PLoS ONE*, vol. 12, no. 6, Joseph, S., Ed., e0179302, 2017. DOI: 10.1371/journal.pone.0179302.

[287] Thor, J. *et al.*, "Application of statistical process control in healthcare improvement: systematic review," *BMJ Quality and Safety in Health Care*, vol. 16, no. 5, pp. 387–399, 2007. DOI: 10.1136/qshc.2006.022194.

[288] Toosi, N. B., Soffianian, A. R., Fakheran, S., Pourmanafi, S., Ginzler, C., and Waser, L. T., "Comparing different classification algorithms for monitoring mangrove cover changes in southern Iran," *Global Ecology and Conservation*, vol. 19, e00662, 2019. DOI: 10.1016/J.GECCO.2019.E00662.

[289] Tortini, R., Mayer, A. L., Hermosilla, T., Coops, N. C., and Wulder, M. A., *Using annual Landsat imagery to identify harvesting over a range of intensities for non-industrial family forests*, 2018. DOI: 10.1016/j.landurbplan.2018.04.012.

[290] Tubiello, F. N. *et al.*, "The Contribution of Agriculture, Forestry and other Land Use activities to Global Warming, 1990-2012," *Global Change Biology*, vol. 21, no. 7, pp. 2655–2660, 2015. DOI: 10.1111/gcb.12865.

[291] United States Geological Survey. "Landsat collection 1," United States Geological Survey. (2019), [Online]. Available: https://www.usgs.gov/land-resources/nli/landsat/landsat-collection-1 (visited on 11/22/2019).

[292] Valderrama-Landeros, L., Santiago, F. Flores-de, Kovacs, J. M., and Flores-Verdugo, F., "An assessment of commonly employed satellite-based remote sensors for mapping mangrove species in Mexico using an NDVI-based classification scheme," *Environmental Monitoring and Assessment*, vol. 190, no. 1, p. 23, 2018. DOI: 10.1007/s10661-017-6399-z.

[293] Valiela, I., Bowen, J. L., and York, J. K., "Mangrove forests: One of the world's threatened major tropical environments," *BioScience*, vol. 51, no. 10, pp. 807–815, 2001. DOI: 10.1641/0006-3568(2001)051[0807:MFOOTW]2.0.CO;2.

[294] Vaswani, A. *et al.*, "Attention is all you need," in *Advances in Neural Information Processing Systems*, vol. 2017-Decem, 2017, pp. 5999–6009. arXiv: 1706.03762.

[295] Verbesselt, J., Hyndman, R., Newnham, G., and Culvenor, D., "Detecting trend and seasonal changes in satellite image time series," *Remote Sensing of Environment*, vol. 114, no. 1, pp. 106–115, 2010. DOI: 10.1016/j.rse.2009.08.014.

[296] Verbesselt, J., Hyndman, R., Zeileis, A., and Culvenor, D., "Phenological change detection while accounting for abrupt and gradual trends in satellite image time series," *Remote Sensing of Environment*, vol. 114, no. 12, pp. 2970–2980, 2010. DOI: 10.1016/j.rse.2010.08.003.

[297] Verbesselt, J., Zeileis, A., and Herold, M., "Near real-time disturbance detection using satellite image time series," *Remote Sensing of Environment*, vol. 123, pp. 98–108, 2012. DOI: 10.1016/j.rse.2012.02.022.

[298] Verbesselt, J., Zeileis, A., and Hyndman, R., *Package 'bfast'*, 2015.

[299] Villamayor, B. M. R., Rollon, R. N., Samson, M. S., Albano, G. M. G., and Primavera, J. H., "Impact of Haiyan on Philippine mangroves: Implications to the fate of the widespread monospecific Rhizophora plantations against strong typhoons," *Ocean and Coastal Management*, vol. 132, pp. 1–14, 2016. DOI: 10.1016/j.ocecoaman.2016.07.011.

[300] Vogelmann, J. E., Xian, G., Homer, C., and Tolk, B., "Monitoring gradual ecosystem change using Landsat time series analyses: Case studies in selected forest and rangeland ecosystems," *Remote Sensing of Environment*, vol. 122, pp. 92–105, 2012. DOI: 10.1016/j.rse.2011.06.027.

[301] Walcker, R. *et al.*, "Fluctuations in the extent of mangroves driven by multidecadal changes in North Atlantic waves," *Journal of Biogeography*, vol. 42, no. 11, pp. 2209–2219, 2015. DOI: 10.1111/jbi.12580.

[302] Walther, G. R. *et al.*, "Ecological responses to recent climate change," *Nature*, vol. 416, no. 6879, pp. 389–395, 2002. DOI: 10.1038/416389a. arXiv: 1109.1006v1.

[303] Wang, C., Guo, H., Zhang, L., Liu, S., Qiu, Y., and Sun, Z., "Assessing phenological change and climatic control of alpine grasslands in the Tibetan Plateau with MODIS time series," *International Journal of Biometeorology*, vol. 59, no. 1, pp. 11–23, 2015. DOI: 10.1007/s00484-014-0817-5.

[304] Wang, H., Long, H., Li, X., Wu, J., and Qiao, Y., "Land condition diagnosis based on multi-resolution analysis and wavelet transform," *2012 IEEE International Geoscience and Remote Sensing Symposium*, pp. 6161–6164, 2012. DOI: 10.1109/IGARSS.2012.6352663.

[305] Wang, L., Jia, M., Yin, D., and Tian, J., "A review of remote sensing for mangrove forests: 1956–2018," *Remote Sensing of Environment*, vol. 231, no. December 2018, 2019. DOI: 10.1016/j.rse.2019.111223.

[306] Ward, R. D., Friess, D. A., Day, R. H., and Mackenzie, R. A., "Impacts of climate change on mangrove ecosystems: a region by region overview," *Ecosystem Health and Sustainability*, vol. 2, no. 4, e01211, 2016. DOI: 10.1002/ehs2.1211.

[307] Watts, L. M. and Laffan, S. W., "Effectiveness of the BFAST algorithm for detecting vegetation response patterns in a semi-arid region," *Remote Sensing of Environment*, vol. 154, no. 1, pp. 234–245, 2014. DOI: 10.1016/j.rse.2014.08.023.

[308] Wei, H., Heilman, P., Qi, J., Nearing, M. A., Gu, Z., and Zhang, Y., "Assessing phenological change in China from 1982 to 2006 using AVHRR imagery," *Frontiers of Earth Science*, vol. 6, no. 3, pp. 227–236, 2012. DOI: 10.1007/s11707-012-0321-3.

[309] Wen, Q., Zhang, Z., Wang, X., Wang, C., and Liu, S., "Classification of Grassland Types by MODIS Time-Series Images in Tibet, China," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 3, no. 3, pp. 404–409, 2010. DOI: 10.1109/JSTARS.2010.2049001.

[310] Werf, G. R. van der *et al.*, "CO2 emissions from forest loss," *Nature Geoscience*, vol. 2, no. 11, pp. 737–738, 2009. DOI: 10.1038/ngeo671. arXiv: arXiv:1011.1669v3.

[311] White, J. C. *et al.*, "Pixel-Based Image Compositing for Large-Area Dense Time Series Applications and Science," *Canadian Journal of Remote Sensing*, vol. 40, pp. 192–212, 2014. DOI: 10.1080/07038992.2014.945827.

[312] White, J. C., Wulder, M. A., Hermosilla, T., Coops, N. C., and Hobart, G. W., "A nationwide annual characterization of 25 years of forest disturbance and recovery for Canada using Landsat time series," *Remote Sensing of Environment*, vol. 194, pp. 303–321, 2017. DOI: 10.1016/j.rse.2017.03.035.

[313] White, J. C. *et al.*, "Confirmation of post-harvest spectral recovery from Landsat time series using measures of forest cover and height derived from airborne laser scanning data," *Remote Sensing of Environment*, vol. 216, pp. 262–275, 2018. DOI: 10.1016/j.rse.2018.07.004.

[314] White, M. A., Thornton, P. E., and Running, S. W., "A continental phenology model for monitoring vegetation responses to interannual climatic variability," *Global Biogeochemical Cycles*, vol. 11, no. 2, pp. 217–234, 1997. DOI: 10.1029/97GB00330.

[315] Wong, C. J. *et al.*, "Estimating Mangrove above-ground biomass loss due to deforestation in Malaysian Northern Borneo between 2000 and 2015 using SRTM and landsat images," *Forests*, vol. 11, no. 9, p. 1018, 2020. DOI: 10.3390/F11091018.

[316] Woodcock, C. E., Loveland, T. R., Herold, M., and Bauer, M. E., "Transitioning from change detection to monitoring with remote sensing: A paradigm shift," *Remote Sensing of Environment*, vol. 238, no. December 2019, p. 111 558, 2020. DOI: 10.1016/j.rse.2019.111558.

[317] Workie, T. G. and Debella, H. J., *Climate change and its effects on vegetation phenology across ecoregions of Ethiopia*, 2018. DOI: 10.1016/j.gecco.2017.e00366.

[318] World Resources Institute and partners. "Global forest watch homepage," Global Forest Watch. (2021), [Online]. Available: https://www.globalforestwatch.org/ (visited on 08/15/2021).

[319] Wulder, M. A., Masek, J. G., Cohen, W. B., Loveland, T. R., and Woodcock, C. E., "Opening the archive: How free data has enabled the science and monitoring promise of Landsat," *Remote Sensing of Environment*, vol. 122, pp. 2–10, 2012. DOI: 10.1016/j.rse.2012.01.010.

[320] Xue, X., Liu, H., Mu, X., and Liu, J., "Trajectory-based detection of urban expansion using Landsat time series," *International Journal of Remote Sensing*, vol. 35, no. 4, pp. 1450–1465, 2014. DOI: 10.1080/01431161.2013.878058.

[321] Yan, Y. E., Ouyang, Z. T., Guo, H. Q., Jin, S. S., and Zhao, B., "Detecting the spatiotemporal changes of tidal flood in the estuarine wetland by using MODIS time series data," *Journal of Hydrology*, vol. 384, pp. 156–163, 2010. DOI: 10.1016/j.jhydrol.2010.01.019.

[322] Yang, Y. C., Ray, P. A., Brown, C. M., Khalil, A. F., and Yu, W. H., "Estimation of flood damage functions for river basin planning: a case study in Bangladesh," *Natural Hazards*, vol. 75, no. 3, pp. 2773–2791, 2014. DOI: 10.1007/s11069-014-1459-y.

[323] Yang, Y., Erskine, P. D., Lechner, A. M., Mulligan, D., Zhang, S., and Wang, Z., "Detecting the dynamics of vegetation disturbance and recovery in surface mining area via Landsat imagery and LandTrendr algorithm," *Journal of Cleaner Production*, vol. 178, pp. 353–362, 2018. DOI: 10.1016/j.jclepro.2018.01.050.

[324] Younes Cárdenas, N., Joyce, K. E., and Maier, S. W., "Monitoring mangrove forests: Are we taking full advantage of technology?" *International Journal of Applied Earth Observation and Geoinformation*, vol. 63, no. April, pp. 1–14, 2017. DOI: 10.1016/j.jag.2017.07.004.

[325] Yu, W. *et al.*, "Spatial-Temporal Prediction of Vegetation Index With Deep Recurrent Neural Networks," *IEEE Geoscience and Remote Sensing Letters*, pp. 1–5, 2021. DOI: 10.1109/LGRS.2021.3064814.

[326] Yuan, F., Sawaya, K. E., Loeffelholz, B. C., and Bauer, M. E., "Land cover classification and change analysis of the Twin Cities (Minnesota) metropolitan area by multitemporal Landsat remote sensing," *Remote Sensing of Environment*, vol. 98, no. 2-3, pp. 317–328, 2005. DOI: 10.1016/j.rse.2005.08.006.

[327] Zabbey, N., Ekpenyong, I. G., Nwipie, G. N., Davies, I. C., and Sam, K., "Effects of fragmented mangroves on macrozoobenthos: a case study of mangrove clearance for powerline right-of-way at Oproama Creek, Niger Delta, Nigeria," *African Journal of Aquatic Science*, 2021. DOI: 10.2989/16085914.2020.1832437.

[328] Zeileis, A., "A unified approach to structural change tests based on ML scores, F statistics, and OLS residuals," *Econometric Reviews*, vol. 24, no. 4, pp. 445–466, 2005. DOI: 10.1080/07474930500406053.

[329] Zeng, H., Jia, G., and Epstein, H., "Recent changes in phenology over the northern high latitudes detected from multi-satellite data," *Environmental Research Letters*, vol. 6, no. 4, 2011. DOI: 10.1088/1748-9326/6/4/045508.

[330] Zewdie, W., Csaplovics, E., and Inostroza, L., "Monitoring ecosystem dynamics in northwestern Ethiopia using NDVI and climate variables to assess long term trends in dryland vegetation variability," *Applied Geography*, vol. 79, pp. 167–178, 2017. DOI: 10.1016/j.apgeog.2016.12.019.

[331] Zhang, A., Lipton, Z. C., Li, M., and Smola, A. J., *Dive into Deep Learning*. 2020, https://d2l.ai.

[332] Zhang, F., Zhu, X., and Liu, D., "Blending MODIS and Landsat images for urban flood mapping," *International Journal of Remote Sensing*, vol. 35, no. 9, pp. 3237–3253, 2014. DOI: 10.1080/01431161.2014.903351.

[333] Zhang, X., Friedl, M. A., and Schaaf, C. B., "Global vegetation phenology from Moderate Resolution Imaging Spectroradiometer (MODIS): Evaluation of global patterns and comparison with in situ measurements," *Journal of Geophysical Research: Biogeosciences*, vol. 111, no. 4, 2006. DOI: 10.1029/2006JG000217.

[334] Zhang, X. *et al.*, "Monitoring vegetation phenology using MODIS," *Remote Sensing of Environment*, vol. 84, pp. 471–475, 2003. DOI: 10.1016/S0034-4257(02)00135-9.

[335] Zhong, L., Hu, L., and Zhou, H., "Deep learning based multi-temporal crop classification," *Remote Sensing of Environment*, vol. 221, no. December 2018, pp. 430–443, 2019. DOI: 10.1016/j.rse.2018.11.032.

[336] Zhou, J., Jia, L., Hu, G., and Menenti, M., "Evaluation of Harmonic Analysis of Time Series (HANTS): Impact of gaps on time series reconstruction," in *Proceedings of the 2nd International Workshop on Earth Observation and Remote Sensing Applications, EORSA 2012*, 2012, pp. 31–35. DOI: 10.1109/EORSA.2012.6261129.

[337] Zhou, J., Jia, L., and Menenti, M., "Reconstruction of global MODIS NDVI time series: Performance of Harmonic ANalysis of Time Series (HANTS)," *Remote Sensing of Environment*, vol. 163, pp. 217–228, 2015. DOI: 10.1016/j.rse.2015.03.018.

[338] Zhu, Z., "Change detection using landsat time series: A review of frequencies, preprocessing, algorithms, and applications," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 130, pp. 370–384, 2017. DOI: 10.1016/j.isprsjprs.2017.06.013.

[339] Zhu, Z. and Qiu, S. "CCDC," Global Environmental Remote Sensing Laboratory. (2021), [Online]. Available: https://github.com/GERSL/CCDC (visited on 12/15/2021).

[340] Zhu, Z. and Woodcock, C. E., "Continuous change detection and classification of land cover using all available Landsat data," *Remote Sensing of Environment*, vol. 144, pp. 152–171, 2014. DOI: 10.1016/j.rse.2014.01.011.

[341] ——, "Object-based cloud and cloud shadow detection in Landsat imagery," *Remote Sensing of Environment*, vol. 118, pp. 83–94, 2012. DOI: 10.1016/j.rse.2011.10.028.

[342] Zhu, Z., Woodcock, C. E., Holden, C., and Yang, Z., "Generating synthetic Landsat images based on all available Landsat data: Predicting Landsat surface reflectance at any given time," *Remote Sensing of Environment*, vol. 162, pp. 67–83, 2015. DOI: 10.1016/j.rse.2015.02.009.

[343] Zhu, Z., Woodcock, C. E., and Olofsson, P., "Continuous monitoring of forest disturbance using all available Landsat imagery," *Remote Sensing of Environment*, vol. 122, pp. 75–91, 2012. DOI: 10.1016/j.rse.2011.10.030.

[344] Zhu, Z. *et al.*, "Continuous monitoring of land disturbance based on Landsat time series," *Remote Sensing of Environment*, vol. In press, 2019. DOI: 10.1016/J.RSE.2019.03.009.

[345] Zhu, Z. *et al.*, "Including land cover change in analysis of greenness trends using all available Landsat 5, 7, and 8 images: A case study from Guangzhou, China (2000–2014)," *Remote Sensing of Environment*, vol. 185, pp. 243–257, 2016. DOI: 10.1016/j.rse.2016.03.036.

[346] van Bijsterveldt, C. E. *et al.*, "How to restore mangroves for greenbelt creation along eroding coasts with abandoned aquaculture ponds," *Estuarine, Coastal and Shelf Science*, vol. 235, p. 106 576, 2020. DOI: https://doi.org/10.1016/j.ecss.2019.106576.

[347] van Oosterzee, Penny and Duke, Norman. “Extreme weather likely behind worst recorded mangrove dieback in northern Australia,” The Conversation. (2017), [Online]. Available: https://theconversation.com/extreme-weather-likely-behind-worst-recorded-mangrove-dieback-in-northern-australia-71880 (visited on 04/16/2022).