Machine learning and remote sensing applications to shoreline dynamics

Martin Samuel James Rogers

Hughes Hall

Department of Geography

University of Cambridge

November 2021

This thesis is submitted for the degree of Doctor of Philosophy

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text. It is not substantially the same as any work that has already been submitted before for any degree or other qualification except as declared in the preface and specified in the text. It does not exceed the prescribed word limit for the Earth Sciences and Geography Degree Committee.

Martin Rogers

November 2021

Abstract

Machine learning and remote sensing applications to shoreline dynamics

Martin Samuel James Rogers

Coastal communities and land covers are vulnerable receptors of erosion, flooding, or both in combination. The accurate, automated, and wide-scale determination of shoreline position, and its migration at the engineering scale (10^{-1} – 10^{2} km), is imperative for future coastal risk adaptation and management. The recent increase in the acquisition and availability of Big Datasets, including multispectral remote sensing imagery, is providing new opportunities to monitor engineering scale rates of shoreline change and other constituents of coastal risk, including changes to human coastal population densities. This increase in data availability comes with novel challenges to devise and utilise methods to store, process, analyse and extract information from these Big Datasets. This thesis assesses the suitability of different Big Data approaches, namely Machine Learning (ML) and non-ML based tools, for the automated extraction of the coastal vegetation edge in remote sensing imagery. Compared to the instantaneous waterline, few vegetation edge methods have been developed and analysis of the coastal zone processes that can be detected using the shoreline proxy remain understudied.

This thesis initially investigates whether non-ML methods are suitable for the extraction of the coastal vegetation edge from multispectral remote sensing imagery. A novel non-ML tool is introduced and applied, CoasTool, which considers the proximity of the instantaneous water line during vegetation edge extraction. CoasTool performance is compared to the outputs derived from well-established threshold contouring techniques and kernel-based methods as well as one form of ML, Support Vector Machines (SVM). Limitations in the performance of these tools, particularly along shorelines with discontinuous or graded vegetation boundaries, provide justification for the application of a separate form of ML, convolutional neural network (CNN), to this task. A novel CNN, VEdge_Detector, is trained and applied to extract the coastal vegetation edge and its outputs are compared to ground-referenced measurements and manually digitised vertical aerial photographs. VEdge_Detector is applied to a time series of

images to detect annual to decadal scale shoreline dynamics discernible using the coastal vegetation edge.

Shoreline change constitutes one element of coastal risk, and this thesis subsequently investigates the viability of integrating multiple ML-derived datasets, pertaining to different aspects of risk, to calculate relative coastal population exposure to shoreline change. The Guiana coastline, northern South America, is one of the most dynamic stretches of coastline in the world and a region where greater than 90% of its population live below 10 m elevation. The identification of locations where coastal populations are at greatest risk to coastal retreat in this region is thus very important to inform coastal risk management decisions. Accordingly, decadal-scale rates of shoreline change calculated using VEdge_Detector derived shoreline positions are combined with secondary, ML-derived, population datasets (WorldPop). The integration of the two ML-based datasets aids the identification of population exposure hotspot locations and discover, previously unpublished, locations where forced migration due to shoreline change has occurred.

In concluding, the relative merits and drawbacks of using ML verses non-ML techniques to detect the coastal vegetation edge are discussed as well as considering the suitability of the coastal vegetation as a proxy of shoreline position. Further discussion is given on the different considerations coastal stakeholders will have when choosing the most suitable tool to use in shoreline detection tasks, including tool performance, speed, transparency, and ease of use. Remaining research gaps and future research requirements are emphasised, including the need for collaboration between different research institutions to suitably train and apply ML tools in the geosciences.

Acknowledgements

Firstly, I would like to thank my lead supervisor, Mike Bithell for the time he has dedicated to me throughout my PhD. He has provided unparalleled guidance and support, particularly when I have been overcoming, sometimes daunting, machine learning problems. I also want to thank my second supervisor, Tom Spencer, for his expert knowledge of coastal systems, which has ensured that I have remained focussed on the coastal zone questions I have been answering. Thank you also to Iris Möller and Geoff Smith for their support and guidance in setting out the original aims for my thesis, and to Sue Brooks for her expert advice on processes and datasets available in Suffolk.

I will never forget the unique and testing period during which I have completed this PhD, brought on by the outbreak of Covid-19 and the associated lockdown. I will forever be thankful to the Cambridge Coastal Research Unit and wider members of the Department of Geography, University of Cambridge, for continuing to provide a sense of community via virtual zoom meetings and coffee breaks.

Thank you finally to my partner, Jamie, who has provided me with unconditional support during the entirety of my PhD. From assuring and believing in me that I could do it, to telling me to ‘make a plan’, your support has been second to none. I look forward to our next challenges together with our newly adopted daughter, Paris.

Funding Acknowledgments

This work was funded through the UKRI NERC/ESRC Data, Risk and Environmental Analytical Methods (DREAM) Centre for Doctorial Training, Grant/Award Number: NE/M009009/1 and is a contribution to UKRI NERC BLUECoast (NE/N015924/1; NE/N015878/1).

Table of contents

List of Figures

Chapter 1:

Chapter 2:

Chapter 3:

List of Tables

Chapter 1. Coastal processes and monitoring

1.1. An introduction to the coastal zone and coastal shoreline change

Coastal zones are located at the interface between land and sea, and are composed of either rocky (Finkl, 2004), sandy (Scott et al., 2011), or muddy (Allen, 2000) substrate, providing habitat for complex ecosystems (Barbier et al., 2011). Their composition includes subtidal, supratidal and intertidal zones, each of which is inundated by water at different frequencies and durations, which are in turn connected to the continental shelf and more inland regions (Swift, 1974). The Low Elevation Coastal Zone (LECZ) is defined as the contiguous areas of land bordering the sea up to 10 m in elevation (McGranahan et al., 2007). The LECZ covers approximately 2% of the Earth's land mass but contains 10-15% of the global human population (Liu et al., 2013; Oppenheimer and Hinkel, 2018). Coastlines provide global benefit; on the oceanic side human civilisations have benefitted from fisheries and aquaculture (FAO, 2015; Blanchard et al., 2017), energy production (Reguero et al., 2015), trade (OECD, 2013), and tourism (Spalding et al., 2017). In adjacent coastal zones, flat topography, fertile land and freshwater sources has enabled agriculture, industry and urbanisation (Neumann et al., 2015; Corine Land Cover, 2018). Coastal habitats also provide a range of ecosystem services, including tourism, recreation, carbon sequestration and educational opportunities (Barbier et al., 2011). Despite their benefit, coastlines are dynamic and changeable.

Meteorological, hydrological, geomorphological, and anthropogenic forces can shape the coastline. Coastal change associated with meteorological events include low pressure weather systems, storms and hurricanes (Plant and Stockdon, 2012; Brooks et al., 2016). Marine hydrological impacts on the coastline are primarily associated with elevated sea levels that can inundate coastal regions, caused by long-term, chronic global mean sea-level rise (SLR) (Rahman et al., 2011; Passeri et al., 2015; Le Cozannet et al., 2019), and shorter-term acute processes, including storm surges (Brooks et al., 2016), wave height and direction (Vitousek et al., 2017) and tides (Pugh, 2004). Geomorphological processes cause coastal change via cliff failure and landslide (Hapke and Plant, 2010), fluvial sediment supply (Allison and Lee, 2004) and cross-shore sediment transport (Jackson et al., 2005; Bergillos et al., 2017). Finally, human activity can influence coastal morphology via construction (Basco, 2006), land reclaimation and artificial sediment abstraction or nourishment (Stronkhorst et al., 2018).

Major inter-annual atmospheric-oceanic circulations drive weather and climate patterns across coastal regions, including the El Niño Southern Oscillation (ENSO) and the North Atlantic Oscillation (NAO) (Hurrell and Deser, 2010; Barnard et al., 2015). ENSO is driven by differences in sea surface temperature and meteorological pressure between the east and west tropical Pacific (Wardlaw et al., 2007; Santoso et al., 2017) and is attributable to ±0.4 m differences in water level elevation (Barnard et al., 2015). Negative Southern Oscillation Index (SOI) years (El-Niño) correspond with warmer waters in the Eastern Pacific, causing stormier conditions in South America and drier conditions in Australia, conversely positive SOI (La-Niña) years commonly correspond to the reverse (Ropelewski et al., 1986). NAO is dependent upon meteorological pressure differences between the subpolar (Iceland) and subtropical (Azores) Atlantic Ocean (Hurrell and Deser, 2010). Strong positive phases of the NAO correspond to increases in precipitation, storminess, wave conditions and mean surface water elevations in the UK and Northern Europe (Olsen et al., 2012). Conversely, colder and drier conditions are commonly observed during negative phases of the NAO (Hurrell and Deser, 2010). In addition to oceanic circulations, there are astronomical cycles that have major implications on tidal height. The most well-known of these events are the diurnal to semi-diurnal cycles (Short, 1991), spring-neap 28-day cycles, and the 18.6 year lunar cycle which globally affects tidal range by an average of 2.2 cm (Baart et al., 2011). These external forcing factors are perturbed by anthropogenic climate change and associated SLR (Slangen et al., 2016).

Anthropogenic climate change and SLR affect coastal storm magnitude and frequency (Committee on Climate Change (CCC), 2018), and NAO and ENSO magnitude and direction (Cai et al., 2015); however, the degree to which anthropogenic actions perturb these processes remains an area of active investigation. Tidal gauges have measured an average increase in sea levels of 1.81 mm yr^{-1} in the last century (Church and Clarke, 2013, p.1138; Wadey et al., 2014); and SLR rates since 1990 are double those since 1900 (Oppenheimer et al., 2019). SLR reduces the ‘freeboard’ between sea levels and coastline elevations, meaning the same magnitude storm event more severely impacts the coastal zone (Haigh et al., 2016). Coastlines with high tidal ranges contain intertidal habitats which are already exposed to large fluctuations in water height, potentially making them more resilient to SLR (Kirwan and Guntenspergen, 2010; Schuerch et al., 2018), but it is unknown where the greatest rates of coastline change rates due to SLR will be. In locations where human shoreline modification structures, such as sea walls, have been constructed, intertidal habitats may be subject to ‘coastal squeeze’, whereby they cannot migrate inland, restricting their ability to respond to SLR (Pontee, 2013). Some coastlines also exhibit a self-regulating response to SLR and other external forcing factors, caused by the presence of feedback loops.

Feedback loops exist between external forcing factors, sediment movement and coastline change dynamics (Cowell and Thom, 1994), which influences how a shoreline responds to the same external forcing factor. Antecedent coastline position, geology and topography influences how external forcing factors move sediment, or cause water inundation (Loureiro et al., 2012; Stokes et al., 2020), producing spatial heterogeneity in coastline response to the same external forcing factors (Loureiro et al., 2012; Robinet et al., 2020). Coastline position and morphology affects the frequency, extent and duration of coastal zone inundation by dissipating wave height and energy (Dongeren et al., 2007; Shepard et al., 2011; Möller et al., 2014), or by facilitating wave height build-up and directional change (Stokes et al., 2020). Coastlines may retain a dynamic equilibrium by accumulating sediment lost during storm events, via cross-shore connectivity, in meteorologically calmer periods (Short and Jackson, 2013). Coastal morphology can moderate or amplify the impacts of both coastal flooding and erosion, highlighting the need to study coastline evolution at a nested range of spatio-temporal scales.

Coastline change can occur instantaneously, driven by individual wave events, through to millennial scale change, caused by fluctuations in mean sea levels driven by glacial- inter-glacial cycles (Cowell and Thom, 1984; Miller and Dean, 2004). The “engineering scale” relates to coastal change at monthly to centennial timescales (10^{-1} – 10^{2} years), along regional to national stretches of coastline ($10^0 – 10^2$ km) (Cowell and Thom, 1994; Figure 1.1). Coastal dynamics at this scale can acutely impact upon coastal zone receptors, including coastal communities, land covers and intertidal habitats (Miller and Dean, 2004), and human activities in the coastal zone can profoundly affect engineering scale coastal dynamics (Cowell and Thom, 1984). The alongshore connectivity of coastlines means engineering-scale dynamics should be monitored both in isolation, and when nested within supra-national to global scale dynamics (Dawson et al., 2009).

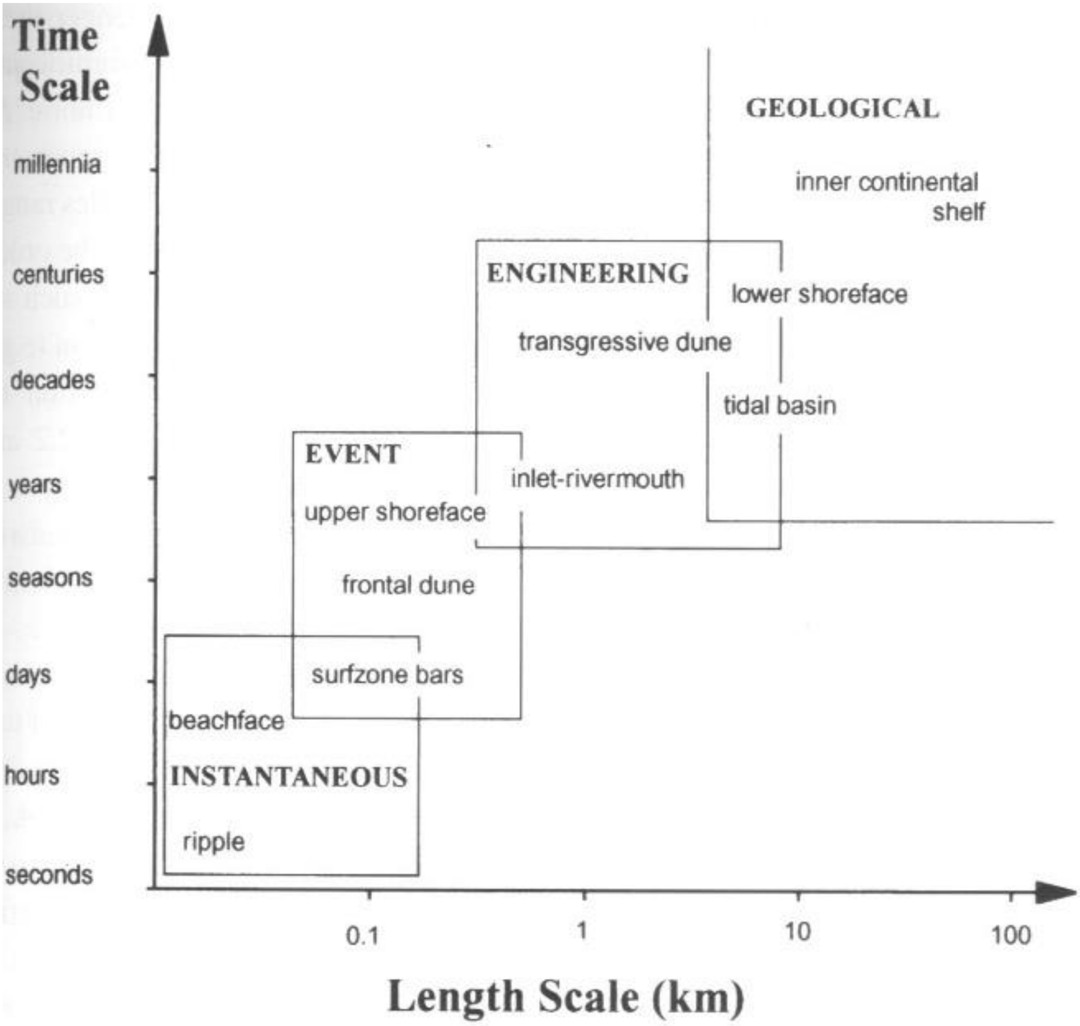


Figure 1.1: The different spatio-temporal scales of coastal morphological change (Cowell and Thom, 1984).

Globally, coastlines are thought to be eroding faster than they are accreting (Pekel et al., 2016; Mentashi et al., 2018; Luijendijk et al., 2018; Vousdoukas et al., 2020), although net rates of change remain disputed (Cooper et al., 2020). The impacts of recent major disasters in the coastal zone highlight their severity to human populations and land covers (CCC, 2018). For example, the 2013 North Sea Storm surge event caused flooding to over 2500 properties in the UK (Environment Agency, 2016), and since the year 2000, hurricanes have caused more than USD $50 billion of damage to coastal regions in the USA (Klotzbach et al., 2018). In coastal zones globally, 4.6% of the world's population are projected to be flooded annually by 2100 without the implementation of measures to reduce these impacts (Hinkle et al., 2014). The strong linkages between shoreline position and coastal flooding and erosion emphasises how coastline change is an important concept assessing the impacts of coastal hazards to coastal receptors.

1.2. Risk in the coastal zone

Coastal risk is defined as the probability of a natural event occurring and the severity of the corresponding impacts to receptors in the coastal zone (Kron, 2013; Rumson et al., 2018). There are three main constituents of risk: hazard, exposure and vulnerability. A hazard is a natural occurrence which has the potential to cause adverse impacts to receptors in the coastal zone (Dawson et al., 2009); but risk only arises when there are vulnerable receptors exposed to it (Kron, 2013). Exposure is the presence and socio-economic 'value' of all receptors which could be adversely impacted by the natural hazard, including population density, businesses, infrastructure, agricultural land and ecological sites (Penning-Rowsell et al., 2014; Calli et al., 2017). Without the presence of human populations or other receptors exposed to the natural hazard, there is no risk (Kron, 2013). Vulnerability relates to physical, socio-economic, political and environmental factors that determine a receptor's susceptibility to the impact of hazards, including health, wealth and mobility (Dwyer et al., 2004; McLaughlin and Cooper, 2010; Bukvic et al., 2020). The two most prevalent and destructive forms of hazard to coastal zone receptors are flooding and erosion (Kron, 2013).

The current global impacts of coastal zone flooding (Dawson et al., 2009), and their rate of increase (Vitousek et al., 2017), are an estimated order of magnitude greater than erosion; but coastline position and dynamics are a key constituent of risk to both coastal flooding and erosion (Section 1.1; Möller et al., 2014). Coastal flood risk schemes, which frequently have not considered changes to coastline position and morphology, have inadvertently exacerbated coastline change rates in adjacent regions by affecting the rate and direction of alongshore sediment movement (Kench, 2013; Nunn et al., 2021); highlighting the interconnected nature between flooding, erosion and coastline position. Section 1.4. outlines the main methods by which shoreline position and its change over time can be monitored.

The impacts of coastline change can only be assessed by concurrently determining the receptors, including human populations, exposed to the hazards. Increases in coastal risk have primarily been attributed to population growth in the coastal zone, instead of an increase in the frequency or intensity of hazards (Pielke et al., 2008; Neumann et al., 2015; Klotzbach et al., 2018). Population growth rates in LECZ's are double those in the hinterlands (McGranahan et al., 2007; Neumann et al., 2015), and coastal populations are projected to increase from over 710 million people in 2015 (Colenbrander et al., 2019) to greater than 1.1 billion by 2100 (Merkens et al., 2016; Brown et al., 2018; Kulp and Strauss, 2019). In some countries, including island states and countries in northern South America, the entire urban population live in the LECZ, reducing the infrastructure available to support any inland migrations forced by coastal hazards (Colenbrander et al., 2019). Measures are required to enable populations in the coastal zone to mitigate and adapt to the impact of coastal hazards, including increased storminess and coastline change.

Factual evidence is required to determine the coastal populations at greatest risk from coastline change (Rumson and Hallett, 2018). Information is needed about a range of disparate processes, from population demographics to historic and current coastline change dynamics. The multifaceted, constantly evolving nature of coastal risk highlights the need for multi-dimensional, high spatio-temporal resolution data to make evidence-based risk-management decisions and inform operational-level emergency response procedures (Meyer et al., 2013; Smith et al., 2017). Datasets pertaining to coastal risk are becoming increasingly publicly accessible, providing new opportunities to gain insight on levels of coastal risk (Rumson and Hallett, 2018; Pollard et al., 2018).

1.3. Big Data in the coastal zone

Big Data refers to datasets containing certain traits, most pertinently high volume, high velocity (data collected and available in near real time) and high variety (containing structured and unstructured data) (Miller and Goodchild, 2015; Kitchen and McArdle, 2016). Although no specific definition of Big Data has been unanimously agreed upon, a range of traits have been devised that distinguish Big Data from other large-volume datasets (Table 1.1). Suitable methods are required to store, pre-process, analyse, visualise and extract knowledge from high-volume, high-velocity coastal datasets (Li et al., 2016; Pollard et al., 2018). This section provides an overview of key datasets pertaining to the two aspects of risk explored within this thesis: shoreline change and coastal zone populations. The approaches employed to extract information and knowledge pertaining to shoreline change and coastal zone populations from Big Data is then explored.

Table 1.1: Key characterises of Big Data, and descriptions of differing levels of these traits. Columns provide a description of a dataset having low, medium or high degrees of each trait. For example, a dataset with low temporal coverage (column 1) may only have entries spanning one month, but a dataset with high temporal coverage (column 5) will have centennial scale records.

	Extent to which dataset contains trait		
Trait	Low	Medium	High
Volume	Megabyte	Petabyte	Exabyte
Velocity	Data available after months - years	Data available after days - weeks	Available instantaneously
Variety	Single data format	Contains > 1 data form e.g. text, numeric, sound, speech	Contains most data forms e.g. text, numeric, sound, speech
Veracity	Noisy, uncalibrated, not quality checked	Quality checked, trustworthy, but containing noise	Quality checked, calibrated, contains little noise
Exhaustivity	Discrete samples	Different locations contain discrete samples or entire coverage	Entire system captured
Scalability	Instantaneous – event scale coverage	Engineering to national-scale coverage	Global coverage
Fine grained (Spatial)	> 10 km resolution	1 km resolution	< 10 m resolution
Fine grained (temporal)	Decadal	Weekly	Hourly - daily
Temporal coverage	< 1 month	1 – 10 years	> 100 years
Relationality	Cannot be combined with other datasets	Pre-processing required to combine	Automated combination with other datasets

1.3.1. Big Datasets pertaining to shoreline position

Datasets used to monitor shoreline position and dynamics can be split into three main categories: satellite remote sensing imagery, unmanned aerial vehicles (UAVs) photographs and ground-referenced measurements (Boak and Turner, 2005). Remote sensing relates to the acquisition of imagery of the Earth's surface using devices not in contact with the target (Boak and Turner, 2005; Toure et al., 2019). Multispectral remote sensing imagery is captured by passive sensors and records the intensity of electromagnetic radiation reflected and emitted from the Earth's surface. Radiation intensity is monitored at discrete wavebands, including red, green, blue (RGB) and near infrared (NIR) (Gao, 1996; McFeeters, 2013). Remote sensing can refer to imagery captured using UAVs and satellite platforms (Toure et al., 2019), but this thesis only refers to satellite imagery when using the term remote sensing. The spatial-temporal resolution and coverage of UAV and remote sensing imagery varies depending upon the platform used to capture the image (Figure 1.2).

Imaging platform	Spatial resolution (m)	Temporal resolution (2021)	Spectral bands
UAVs (Various)	0.12 - 0.2*	Tasked*	3/4*
PlanetScope (Planet)	3	Daily – weekly	4
RapidEye (Planet)	5	Weekly**	5
Sentinel II (European Union)	10	5 days	13
Landsat 5 – 8 (NASA)	30	14 days	9

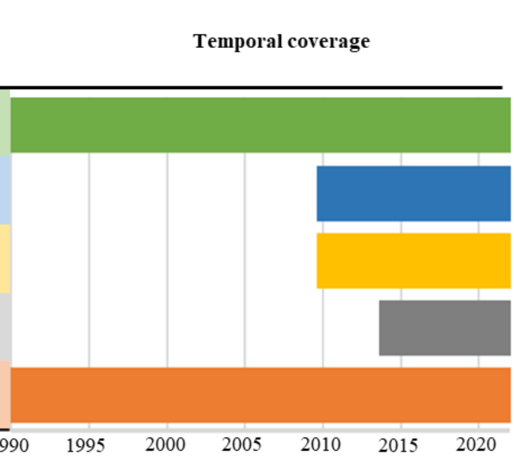


Figure 1.2: Comparison of the different properties of multispectral imagery captured from UAVs and the four satellite platforms used in this thesis: PlanetScope, RapidEye, Sentinel II and Landsat 5 – 8 (Marta, 2018; European Space Agency (ESA), 2021; United States Geological Survey (USGS), 2021). *: the number of spectral bands and spatio-temporal resolution of imagery captured using UAVs will depend upon the UAV used and the frequency of data collection. ** RapidEye was decommissioned in 2020, but images were captured approximately every week until this date.

One of the world's largest ground-referenced datasets of coastline position is the repository of cross-shore elevation profiles (EP) collected by the Environment Agency, which is an executive, non-departmental public body, sponsored by the Department for Environment, Food and Rural Affairs (Defra) (Environment Agency, 2010; Environment Agency, n.d). This monitoring programme has collected high veracity (± 10 mm positional accuracy), biannual profiles along the East Coast of England at 0.5 km transect intervals since 1991 (Environment Agency, 2010; Environment Agency, n.d). The EP dataset, therefore, contains high volume, veracity and temporal coverage, but contains other Big Data traits to a lesser degree than remote sensing imagery (Table 1.1).

Some Big Data traits contained more within remote sensing imagery than EP, which are useful in engineering-scale studies of coastline change, are exhaustivity, scalability and velocity. The EP dataset has low exhaustivity, because it contains samples from discrete locations, rather than capturing an entire system, meaning shoreline positions must be interpolated between these discrete monitoring locations. Remote sensing provides more consistent spatial coverage of a study area (Cenci et al., 2018). The low scalability in EP data exists because data is only located within a finite location (Humber to Thames Estuary, England) (Environment Agency, n.d.), meaning coastal dynamics outside of the study area are not captured. Except for high latitude regions, most satellite platforms provide imagery of the entire globe, including locations which are inaccessible or dangerous to get to (Gorelick et al., 2017). EP samples contain low velocity because they are only collected every six months. This precludes the use of the dataset to, for example, determine the impact of an individual storm event, unless data was fortuitously collected in the immediate aftermath of the event. Some satellite platforms now collect daily remote sensing imagery, which is made publicly available within 24 hours (Marta, 2018), making it suitable to investigate the impact of storm events (Splinter et al., 2018), if suitably cloud-free imagery is available. As such, whilst the use of remote sensing imagery should never replace the collection of ground-referenced measurements of shoreline position, satellite imagery provides the best opportunity to monitor shoreline position and change at the engineering scale. It is necessary, therefore, to explore the range of methods available to extract the shoreline position from remote sensing imagery.

1.3.2. Big Data applications to shoreline change

Methods are required to extract engineering to global scale shoreline position from remote sensing imagery. The coastlines of smaller ($10^0 – 10^1$ km) study areas have been manually digitised (Ferreira et al., 2006; Theiler et al., 2013) but these methods are unpractical and time-consuming when applied to Big Datasets consisting of coastlines longer than 10 km, or to multiple short stretches of coastline, meaning automated techniques are required. Before applying automated shoreline detection techniques, it is necessary to determine which feature will be extracted to represent coastline position (Toure et al., 2019).

A shoreline proxy is a visibly discernible feature in multispectral remote sensing imagery (Boak and Turner, 2005). Shoreline proxies can be broadly classified into coastal zone geomorphological, vegetation, water or human features (Toure et al., 2019; Pollard et al., 2020). The instantaneous waterline position is the dominant shoreline proxy extracted from optical remote sensing imagery (Boak and Turner, 2005), because it is present in all coastal zones and closely relates to the coastline as the land-water interface (Vos et al., 2019a).

However, collating a time-series of instantaneous water line position in isolation does not necessarily provide an indication of engineering-scale net shoreline migration. The amplitude of horizontal change in waterline position caused by diurnal or semi-diurnal tidal cycles can vary depending on beach gradient, which in turn is often linked to beach sediment size and sorting (McLean and Kirk, 1969; Komar, 1998). So, depending on where in the tidal frame the image was captured, tidal range potentially has a greater effect on waterline position than decadal shoreline accretion or erosion (Pugh and Woodworth, 2014). Crucially, this means large-scale changes in waterline position may have no bearing on the level of risk coastal population face to coastline change, erosion and flooding.

Methods have been devised to try and overcome these limitations in detecting the waterline position from remote sensing imagery. The mean waterline position has been extracted from multiple, temporally adjacent, images (Almonacid-Caballer et al., 2016) but this removes the ability to detect short-term variability and, even then, there are spring-neap, equinoctial and nodal tide cycles operating at different timescales. Waterline position can be tidally corrected by considering slope profile and tidal stage during image capture (Vos et al., 2019a); although approximate slope profiles are required when concurrent datum-based measurements are not available. Thus, given the difficulties of deriving a robust waterline position indicator, there is potential value in seeking out alternative shoreline proxies from remote sensing imagery to quantify temporal rates of shoreline change.

The vegetation line represents the most seaward extent of plant species and communities (Allen, 2000; Miller et al., 2010). It can be flood-responsive, representing the limits to spring high tide flooding, or erosion-responsive, delineating the boundary between the upper beach and the base of sand dunes or soft rock cliffs (Pollard et al., 2019b; Toure et al., 2019). The proxy remains heavily understudied compared with the waterline (Toure et al., 2019), because it commonly forms a discontinuous and heterogeneous boundary. The vegetation line, however, is more stable than the waterline, and can provide insight into backshore dynamics, which are processes associated with extreme wave and tide events, and which cannot be detected by focussing solely on the land-water interface (Grzegorzewski et al., 2011; Toure et al., 2019). Vegetated coastal ecosystems, including sand dunes and salt marshes, commonly act as the first line of defence for landward coastal populations (Grzegorzewski et al., 2011; Wagner et al., 2017). Vegetation line position and dynamics are therefore intrinsically linked to the current and future risk of populations to coastal hazards which justifies the development of new automated methods for its detection.

Automated methods used to extract the waterline, which have yet to be applied to vegetation edge detection include threshold contouring, classification of land covers and kernel-based operators. Threshold contouring applies a discrete value to a spectral band, or combination of bands, and contours at this value to estimate coastline position (McFeeters, 1996; Hagenaars et al., 2018). Land cover classification clusters remote sensing image pixels into discrete groups, corresponding to different land cover classes, including water, vegetation, sand and urban categories (Pekel et al., 2016). Kernel-based operators pass a filter, commonly a 3×3 grid, over the image to identify locations with the greatest gradient change in spectral values (Pardo-Pascual et al., 2012). Classification and threshold contouring are discussed in more detail in Chapter 3, and kernel-based methods are described in Chapters 2 and 3. There is a need to determine the performance of these well-established tools against other newly emerging tools, namely machine learning.

Machine learning (ML) is a Big Data approach that identifies patterns and relationships in datasets (Goodfellow et al., 2016). ML tools are inductive and determine relationships between inputs and outputs to derive predictions (Goldstein et al., 2019; Kim et al., 2019). This inductive reasoning distinguishes ML tools from the above-mentioned methods that require human input to manually define rules (Jordan and Mitchell, 2015). For example, the threshold-value in threshold contouring (Vos et al., 2019b), and the number of classes in image classification (Wickham et al., 2013), both require inductive reasoning. ML may overcome the limitations of other tools where the heterogeneity of the spectral properties of coastal features, forces manual rules in non-ML tools to be iteratively updated (Liu et al., 2019). This could frustrate the automated nature of non-ML tools detecting the coastal vegetation edge, because the spectral properties of the coastal vegetation will vary substantially due to differences in vegetation species, phenology, composition and density (Unberath et al., 2019). However, despite their perceived advantages, ML tools have been criticised for being difficult to train and interpret (Rudin, 2019).

ML tools are often considered to be 'black box' tools, where it is not easy to determine how they derive their outputs (Rudin, 2019). Inconsistencies in ML tool performance may also be difficult to explain, where a ML tool accurately predicts outputs from one dataset but cannot generalise to others (Jabbar and Khan, 2015). An ML tool is considered to be ‘fragile’ when the output predictions for two separate, but similar, images varies greatly (Jabbar and Khan, 2015). In edge detection this could lead to, for example, a trained ML tool being able detect shoreline position in one set of remote sensing images but not in other images. Poor or inconsistent performance in ML is commonly attributed to the tool not being exposed to a sufficiently large dataset to learn patterns (Saravanan et al., 2018). It is unknown whether there is a sufficient quantity or variety of remote sensing data available to train a ML tool to automatically detect the coastal vegetation edge, or whether the volume of imagery required is too large to make the task feasible. Further investigation will help determine whether ML or non-ML tools are more suitable for the automatic detection of the coastal vegetation edge in remote sensing imagery. ML tools and their applications to coastal risk are described in more detail in Chapter 2.

1.3.3. Big Data applications to human dynamics in the coastal zone

Alongside shoreline change dynamics, there is a need to identify the size and distribution of the population exposed to this hazard. Identifying coastal populations at greatest risk to shoreline change is necessary because financial, infrastructural, technical, and physical barriers prevent the implementation of many mitigation and adaptation schemes that would reduce the impacts of shoreline change to coastal populations (Hinkel et al., 2014). A limited source of funding poses the main barrier to the implementation of schemes to reduce the impacts of shoreline to coastal communities (Hinkel et al., 2018), particularly when stakeholders disagree on who should pay (Aerts et al., 2014), which receptors are the most important to protect (Barquet and Cumiskey, 2018) or whether the costs of a scheme outweighs the benefits (Hinkel et al., 2014; Penning-Rowsell et al., 2014; Hinkel et al., 2015). These barriers highlight the importance of identifying suitable datasets and Big Data approaches to that can identify populations at greatest risk to coastline change to target risk reduction schemes.

Two datasets providing information on coastal zone population are the census (e.g. Office for National Statistics (ONS), 2021), and auxiliary data e.g. night light and land cover maps, which provide a proxy of population density (Stevens et al., 2015). The census is a well-cited example of a large-volume dataset of population dynamics which does not contain many other Big Data characteristics Kitchin and Lauriault, 2015). The UK census dates back to 1801, providing unrivalled temporal coverage (ONS, 2021), and high exhaustivity of the UK population system (Kitchin and Lauriault, 2015); but census data is only collected every ten years in the UK (low velocity), is inconsistently collected between nations (low relationality), and is spatially aggregated when made publicly available (coarse spatial resolution) (Fotheringham and Wong, 1991). The lack of some Big Data traits within census data has previously limited its use in coastal risk studies, for example it could not be used to detect localised human landwards migration caused by coastal hazards (Hauer et al., 2019), prompting the use of other population datasets in coastal risk studies.

Datasets which could provide a proxy of population density include night light maps, road and infrastructure network maps, and multispectral remote sensing imagery classified into different land cover (Stevens et al., 2015; Tatem et al., 2018). Machine learning based methods available to generate gridded population density maps from these auxiliary datasets are detailed in Section 2.7.1. Compared to the census, these auxiliary datasets generate higher velocity, higher variety and finer spatio-temporal resolution population data; but they contain lower temporal coverage and do not capture the entire population: for example, informal settlements that have little permanent street lighting (low veracity) (Wang et al., 2019). The high veracity of census data means it remains a robust information source in locations where population growth has remained stable (Jäger et al., 2018); however, where census data is missing, or an area experiences rapid population accumulation, other datasets containing Big Data traits may be more suitable.

In summary Big Datasets such as satellite remote sensing imagery and auxiliary datasets are opening new opportunities to monitor changes in shoreline position and coastal zone populations with higher spatial-temporal frequency and coverage. For example, terabytes of remote sensing imagery is collected every day, enabling regular monitoring of shoreline dynamics and coastal zone populations over spatial scales for which it is unfeasible to collect ground-referenced measurements (Gorelick et al., 2017; Tamimina et al., 2020). The high velocity of Big Datasets also provides promise in detecting the impacts of coastal hazards or forced human migrations in real-time, which could be very informative for affected communities and risk-management authorities (Pollard et al., 2018). Appropriate methods are required to extract knowledge and information from these Big Datasets. Machine learning based approaches have been shown to provide insight from Big Data in a range of disciplines spanning medicine and navigation (Jordan and Mitchell, 2015), yet applications into coastal zone dynamics remain relatively under explored. Further investigation is required to determine the additional knowledge, pertaining to coastal risk, which can be gleaned using Big Datasets and Big Data approaches such as machine learning.

1.4. Thesis outline

This thesis assesses the suitability of different ML and non-ML based tools for the automated extraction of the coastal vegetation edge in remote sensing imagery. The viability of integrating multiple ML-derived datasets, pertaining to different aspects of risk, is investigated.

Chapter 2 provides a description of the key forms of ML used in this thesis. It reviews current applications of ML to different aspects of coastal risk, including coastal hazards, namely flooding and erosion, receptor vulnerability and exposure. ML applications to coastal risk are in their infancy, and this chapter highlights persisting research gaps which are explored in the rest of the thesis.

The framework for data Chapters 3, 4 and 5 is visualised in Figure 1.3. Chapter 3 investigates whether non-ML methods are suitable for the extraction of the coastal vegetation edge from multispectral remote sensing imagery. A novel, non-ML tool is introduced and applied, CoasTool, which considers the proximity of the instantaneous water line during vegetation edge extraction. The tool is applied to four locations across the United Kingdom, representing different forms of intertidal habitat and morphology. These results are compared to outputs derived from well-established threshold contouring techniques, kernel-based methods and one form of ML, Support Vector Machines (SVM). The first research question thus investigates the relative performance of these different ML and non-ML tools in identifying the position of the coastal vegetation edge via remote sensing imagery.

The results of Chapter 3 provide justification for further training and application of a separate form of ML, convolutional neural network (CNN), to this task. Chapter 4 trains and applies a novel CNN, VEdge_Detector, for the automated extraction of the coastal vegetation edge. A dataset containing greater than 30,000 images is generated to train VEdge_Detector. To assess performance, VEdge_Detector outputs are compared to ground-referenced measurements and manually digitised vertical aerial photographs across seven test sites. VEdge_Detector is used to identify the vegetation edge in satellite images of soft rock cliff margins at Covehithe, Suffolk, U.K, captured over multiple years, to detect annual to decadal scale shoreline dynamics. Shoreline change results are discussed in the context of recent major North Sea storm events. To further address research question 1, the performance of VEdge_Detector is compared to that of all ML and non-ML tools outlined in Chapter 3. Research question 2 specifically determines whether it is possible to train and apply a CNN to the task of automated vegetation edge detection, and whether CNN outputs can be used to determine rates of shoreline change.

Chapter 5 upscales VEdge_Detector by applying it to the entire Guiana coastline, northern South America, spanning three countries and a coastline length greater than 1500 km. The Chapter beings by justifying why the Guiana coastline is studied, namely that greater than 90% of the total population, or approximately 1.5 million people, in the Guianas live in the LECZ, and that the Guiana coastline is home to one of the most dynamic coastlines globally (Mentashi et al., 2018; Colenbrander et al., 2019). The vegetation edge is extracted from imagery captured between 1990 and 2021, determining annual to decadal-scale trends in shoreline change. Shoreline dynamics are statistically correlated against the North Atlantic Oscillation index, the El-Niño Southern Oscillation index and the 18.6-year nodal cycle. Total populations living in close proximity to the Guiana coastline are calculated using secondary gridded population datasets: WorldPop (Stevens et al., 2015; Section 2.7.1). These population and shoreline change datasets are combined to identify exposure hotspot locations and discover, previously unpublished, locations where forced migration due to shoreline change has occurred. Research question 3, accordingly, asks whether a CNN can be applied to detect shoreline change at a supra-national scale. It investigates what coastal dynamics can be gleaned using supra-national scale studies, compared with using local-scale shoreline change datasets, and determines whether different datasets can be integrated to ascertain relative populations exposed to shoreline change,

Chapter 6 summarises and synthesises the findings of the thesis. It emphasises remaining research gaps and future research requirements, particularly in relation to the training and application of ML tools in coastal risk studies. The need to improve the transparency of ML methods to increase uptake is also discussed.

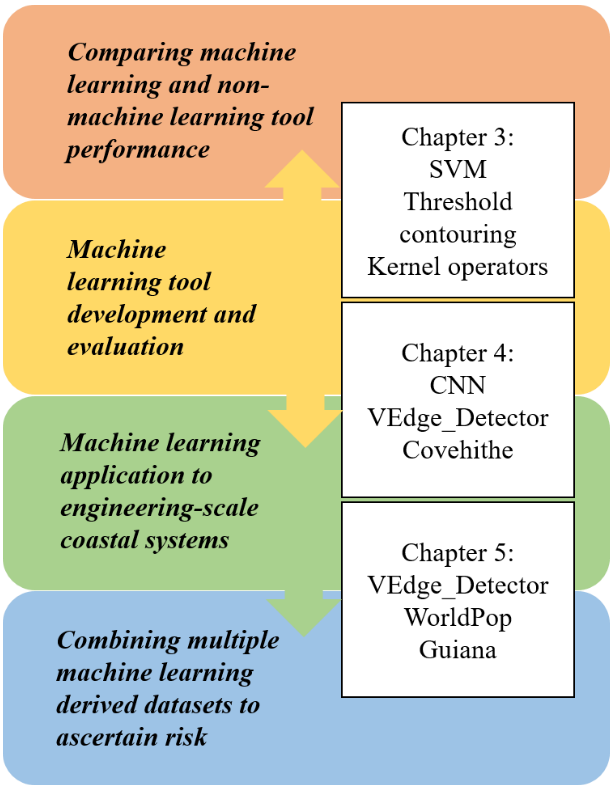


Figure 1.3: Schematic of the structure of this thesis.

1.5. Research questions and objectives

Aim

To evaluate the extent to which machine learning (ML) and remote sensing tools can be used to improve understanding of the different levels of risk to coastal populations arising from shoreline change.

Research questions and objectives

1. Which techniques can be used to automatically detect the coastal vegetation edge via multispectral remote sensing imagery?
    a. To review the range of ML and non-ML based techniques that could be used to extract the coastal vegetation edge in multispectral remote sensing imagery.
    b. Compare the performance of a range of pre-existing and novel non-ML techniques for detecting the coastal vegetation edge.
    c. Evaluate the performance of non-ML tools and a ML-based pixel classifier, support vector machines, at detecting the coastal vegetation edge across a range of intertidal habitats.

2. Can ML techniques, namely convolutional neural networks, be used to detect the coastal vegetation edge and rates of change using remote sensing imagery?
    a. Review the differences in using an object-based classifier, such as convolutional neural networks, to detect the coastal vegetation edge compared with pixel-based classifiers.
    b. Train a convolutional neural network to automatically extract the coastal vegetation line from multispectral imagery, considering the impact of input image spectral band selection on tool performance.
    c. Calculate decadal-scale rates of water and vegetation line change. Identify different coastal zone processes that can be detected using the different shoreline proxies.

   d. Examine the differences in the expertise and resource required to build, train and apply different ML and non-ML tools to coastal vegetation edge detection.

3. Can multiple ML tools be integrated to estimate relative levels of coastal populations exposed to shoreline change at a supra-national scale?
   a. Use convolutional neural networks to extract the coastal vegetation edge and calculate decadal, supra-national scale rates of shoreline change across the Guiana Coastline, northern South America.
   b. Review the range of Big Datasets available to calculate supra-national scale patterns in coastal zone populations.
   c. Examine the suitability of combining ML-based datasets pertaining to shoreline change and coastal population densities to identify areas of high population density lying adjacent to rapidly retreating stretches of coastline.

Chapter 2. Machine learning applications to coastal risk

Chapter 1 summaries examples of pre-existing tools, which do not use machine learning (ML) techniques, to extract the position of the shoreline from multispectral remote sensing imagery. These non-ML based tools that could be applied to detect edges in remote sensing imagery, including threshold contouring and kernel-based operators, are described in more detail in Chapter 3. ML tools are a suite of methods, separate to non-ML tools, whose performance at detecting the coastal vegetation edge remains unstudied. Prior to applying ML tools to this task, it is necessary to define ML, distinguish between the different forms of ML, and review the current applications of ML to coastal risk.

2.1. What is machine learning?

Machine learning (ML) is a form of artificial intelligence that uses data to train a computer system to learn patterns and processes (Goodfellow et al.,2016). ML tools improve their performance through experience, whereas other forms of statistical and process-based modelling rely on humans to manually define key terms and values (Jordan and Mitchell, 2015). This point can be illustrated using one prominent example of ML: driverless cars. As a computer system is provided with more images of road scenes manually labelled by humans, (e.g. “car”, “pedestrian” and “lamppost”), the system can increase its ability to automatically identify these features in new, unlabelled images; reducing the need for human intervention (Janai et al., 2020). ML tools have been applied to many domains, including computer vision, finance, healthcare and astronomy (Jordan and Mitchell, 2015; Miotto et al., 2018; Carleo et al., 2019), but applications to coastal risk remain in their infancy (Goldstein et al., 2019).

There are multiple forms of ML, but this thesis uses two types: convolutional neural networks (CNN) and support vector machines (SVM). These types of ML are used alongside secondary population data derived using a separate form of ML, random forests (RF) (Tatem et al., 2017). These tools have been used in various data analytical tasks, including image classification, pattern recognition, identifying correlations and causality between variables and parameter predictions; LeCun et al., (2015) and Goldfellow et al., (2016) provide an overview of ML functions. The mathematical formulas underpinning these ML tools have been established for many decades (LeCun et al., 1989; Corpes and Vapnik, 1995; Breiman et al., 2001; Wang and Raj, 2017) but recent advances in computer processing power, and the availability of real-time environmental data, is opening new opportunities to use these ML tools. Packages have been developed in R, Python and other programming languages enabling researchers to train and develop ML tools in a few hundred lines of code, thus making them more accessible to different research domains (Abadi et al., 2016). After discussing key principles relevant to all ML tools, this section describes CNN, SVM, and RF. Subsequent sections will then review applications of these ML tools to different aspects of coastal risk, with a focus on coastline change and population dynamics. Another ML technique, Bayesian Networks (BN), which allows the developer to include prior knowledge on the relationship between particular parameters, is not used in this thesis but studies that have used BN are also contained within this chapter.

2.2. Principles pertinent to multiple forms of machine learning

There are two main sub-categories of ML: supervised and unsupervised learning. Supervised ML tools are provided with both the input and corresponding output datasets (Goldstein et al., 2019) whereas unsupervised ML tools are only provided with the input data (Alloghani et al., 2020). Unsupervised ML tools are primarily used to classify data into subgroups, reduce dimensionality in the input dataset, or identify underlying patterns or signals in the data (Gentlemen and Carey, 2007). Due to the availability of ground-based and remote sensing derived output data on shoreline position, this thesis exclusively uses supervised forms of CNN, SVM and RF.

Supervised ML tools are trained to determine formula weights which produce outputs that adequately reach the output data values when provided with the input data (Glorot and Bengio, 2010; Goodfellow et al., 2016). In its simplest form, linear regression is an example of supervised ML. When provided with a dataset containing input and output values, the computer determines the line gradient, m, and intercept, c, which minimises the difference, or error, between the predicted and observed output values. Supervised ML training consists of three stages: training, validation and testing. During the training stage, a large proportion of the available dataset, commonly 70 – 80%, is used to iteratively update the function weights and bias (analogous to m and c), in order to reduce the error, or loss, between the predicted and observed output data values (Glorot and Bengio, 2010). During validation, predictions are made by the trained tool on some 20% of the remaining data, previously unseen by the tool. The errors between predicted and observed output data values are calculated and used to further update the formula weights (Larochelle et al., 2009). This validation stage checks for overfitting- where a ML tool is very capable at deriving the output from the input values in the training dataset but is not able to generalise well to previously unseen data entries. A ML tool overfits the training dataset, and is said to have high variance, if its performance substantially differs when applied to different datasets (Geman et al., 1992). During testing, tool performance is assessed on the final 5 – 10% of the data. This final stage checks for bias within the validation dataset, i.e. the validation dataset may have inadequately represented the range of values contained within the entire dataset (underfitting) (Dietterich and Kong, 1995). The aim of supervised ML training, validation and testing is to reduce both the bias and variance contained within the ML tool (Larochelle et al., 2009).

Using small training datasets can lead to the ML tool overfitting because it has not been exposed to a range of possible variable values and combinations (Larochelle et al., 2009). In this circumstance, data augmentation methods can be used to artificially increase the size of the dataset. Augmentation methods include bootstrapping, whereby the ML tool is trained multiple times on a different, randomly generated, sample of the training dataset (Efron and Tibshirani, 1994). The sample is commonly the same size as the original dataset but there may be duplicate entries within the subset, enabling a ML tool to be trained on a different combination of data entries multiple times (Efron and Tibshirani, 1994). In image-based ML tasks, augmentation can including flipping, rotating or subsetting the image (Perez and Wang, 2017). These methods can reduce variance and bias in trained ML tools when an otherwise small training dataset would be available (Perez and Wang, 2017).

The time-consuming nature of training a supervised ML tool, including deriving a training dataset via data augmentation techniques, justifies an initial investigation into the performance of non-ML methods in automatically detecting the coastal vegetation edge. A potential benefit of some ML methods, including SVM and RF, is that they can be trained using smaller training datasets compared with CNNs (Elgohary et al. 2017). The recent advances in computer vision accomplished using CNNs, including object detection in visible imagery (Janai et al., 2020), justifies investigation into their ability to detect edges in remote sensing imagery if other ML and non-ML tools cannot achieve this task with adequate precision and accuracy. The following sections accordingly discusses some of the key principles of SVM, RF and CNN in turn.

2.3. Support Vector Machines

Support Vector Machines (SVM) categorise data entries into two or more categories, and have commonly been used to, for example, classify pixels in multispectral remote sensing imagery into different land cover categories (Maxwell et al., 2018; Mountrakis et al., 2011). SVM derive a hyperplane which separates the data entries into discrete classes in feature space (Figure 2.3 (a); Cortes and Vapnik, 1995). The 'support vectors' are the individual data entries in each class which lie closest to the separating hyperplane (Maxwell et al., 2018). During training, SVM aims to minimise classification error and maximise the distance between the support vectors and the hyperplane (Elnabwy et al., 2020). After each iteration of SVM model training, the misclassification errors and distances between the hyperplane and the support vectors are calculated and this informs the model on how to alter the position of the hyperplane during the next iteration (Mountakis et al., 2011).

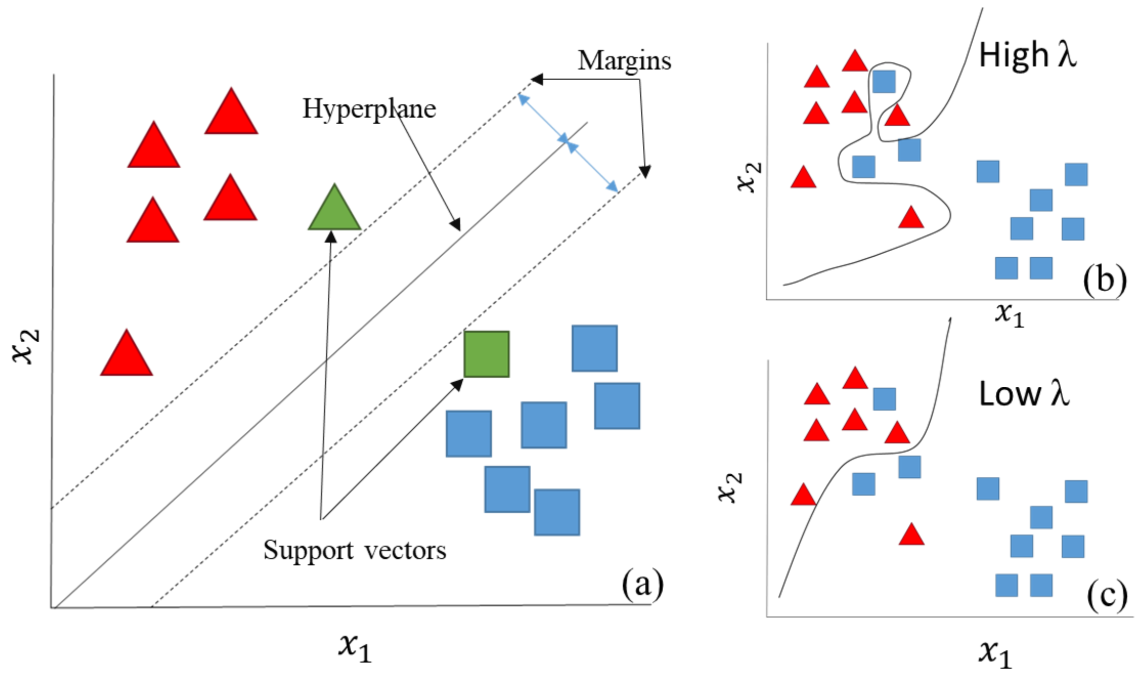


Figure 2.1: Key features pertaining to SVM. (a) Simple example of how SVM produce a hyperplane separating all data entries into two categories. Comparison of hyperplane produced when using a (b) high and (c) low regularisation parameter (λ) value.

A number of hyperparameters, most notably the regularisation parameter (λ) and kernel function, alter SVM performance and may be manually tuned prior to SVM training. A larger λ value increases the penalty applied when a data entry is misclassified by the model (Figure 2.1 (b) – (c); Maulik and Chakroborty, 2017). A higher λ value results in fewer misclassified pixels in the training dataset, but if λ is set too high the model is liable to overfit the training dataset, reducing performance when applied to a testing dataset (Figure 2.1 (b)). By contrast, a low λ may produce a hyperplane which is too smooth, and the model may misclassify a high proportion of data entries, resulting in high model bias (Figure 2.1 (c); Mountrakis et al., 2011; Zhang et al., 2013). Establishing the most suitable λ is commonly derived from testing a number of candidate values and choosing the value with the lowest error term (Maulik and Chakroborty, 2017). SVM can generate linear hyperplanes, or kernels, including polynomial kernels, can be used to generate non-linear hyperplanes (Suthaharan, 2016; Choung and Jo, 2017; Maulik and Chakroborty, 2017). Suthaharan (2016) provides a contemporary summary of SVM theory and applications.

SVMs are a form of ML that can be used to classify pixels in remote sensing imagery. They provide promise in classifying and detecting features in remote sensing imagery because they can process high dimensional datasets and outperform other classifiers, such as Maximum Likelihood Estimator (MLE), when trained using a small training dataset (Elgohary et al., 2017). Unlike MLE, SVM are non-parametric and do not assume the training dataset is normally distributed. This is pertinent for satellite imagery which commonly contains high levels of noise (Maulik and Chakroborty, 2017). The training data requirements of SVM provide promise in the method being able to detect the coastal vegetation edges in remote sensing imagery, and accordingly justifies investigation into their ability to conduct this task.

2.4. Random Forests

Random Forests (RF) consist of an ensemble of decision trees which individually split a dataset into increasingly homogenous categories (Genuer et al., 2017). Decision trees, also known as Classification And Regression Trees (CARTs), split the dataset multiple times into smaller sub-classes using threshold values (Figure 2.2 (a); Bayram et al., 2017; Cutler et al., 2012). Bootstrapping is commonly applied so that each decision tree is trained on a different random sample of the original input dataset (Figure 2.2 (b)); Breiman et al., 2001). During training, a decision tree starts with all data at the root node, and the tree 'grows' by splitting the data via decision nodes to terminal nodes (Figure 2.2 (b); Breiman et al., 2001). The decision trees are independently grown, meaning each tree has a different architecture and data entries in their root nodes.

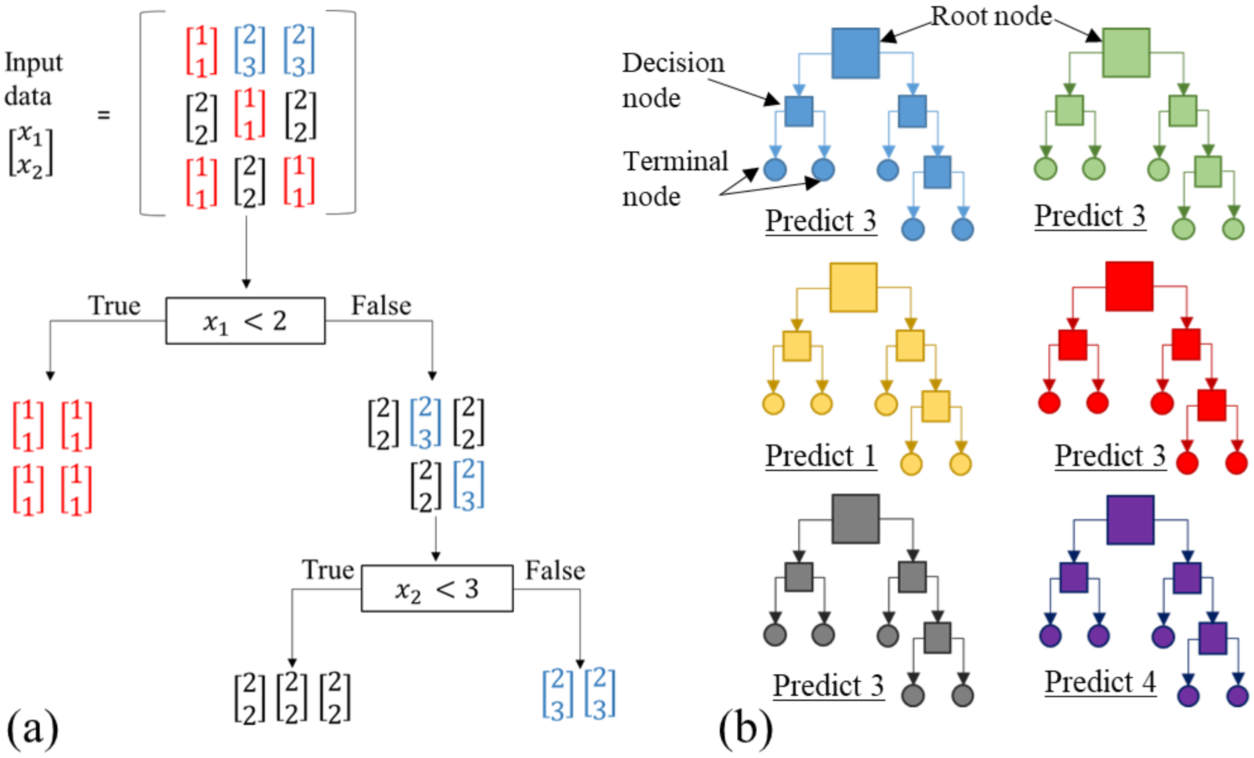


Figure 2.2: Overview of Random Forests (RF) architecture.(a) Schematic of an individual decision tree. (b) Ensemble of decision trees combined to produce a RF. In this example, output class 3 is the most commonly predicted class by the decision trees, so this class forms the RF prediction.

The trained decision trees contained within a RF individually assign an output class to a previously unseen data entry. Where the output class predicted by the decision trees is not unanimous, the output class predicted by the most decision trees is assigned as the RF predicted class (Figure 2.2 (b); Genuer et al., 2017). If a single decision tree is used, there is a risk of high variance and overfitting in the model. This is because the tree is only capable of splitting the data that it has been provided with. Because the ensemble of decision trees in the RF have all been trained on a different dataset, the likelihood of overfitting in RF predictions is reduced (Breiman et al., 2001). See Breiman et al., (2001) for more detailed descriptions of RF architecture, training and functions.

This thesis uses third party WorldPop data, generated using RF, which provides population dynamics at 100 m spatial resolution (Tatem et al., 2017). These datasets have been selected because RFs are non-parametric prediction models, meaning they do not assume any predetermined relationship between the input predictor variables and the dependent variable (Breiman et al., 2001; Stevens et al., 2015). RFs are also able to conduct analysis using many input predictor variables. This ability is advantageous when using multiple remote sensing datasets, with variable resolutions, in different coastal locations. A separate form of ML, Convolutional Neural Networks, are also well suited to the analysis of multi-dimensional remote sensing data.

2.5. Convolutional Neural Networks

Convolutional Neural Networks (CNNs) have larger training requirements that SVMs and RFs, but their ability to derive semantic information via convolution, as discussed below, provides promise in them being able to detect features in remote sensing imagery. CNNs have been applied to remote sensing images to detect features, extract edges and conduct pixel-based classification (Kattenborn et al., 2021). The architecture of a CNN is constituted of its input layer, hidden layer(s) and output layer. Each layer contains a different number of nodes and synapses connect the nodes in one layer to nodes in adjacent layers (Figure 2.3 (a)). Deep learning refers to CNNs which contain two or more hidden layers (Chen et al., 2016). Where a synapse links two nodes together, its weight, w, corresponds to the number to multiply the node value by, to get the value for the node in next layer (Figure 2.3 (b)). The activation function, σ, enables the CNN to determine non-linear relationships between input and output variables (Ramachandran et al., 2017).

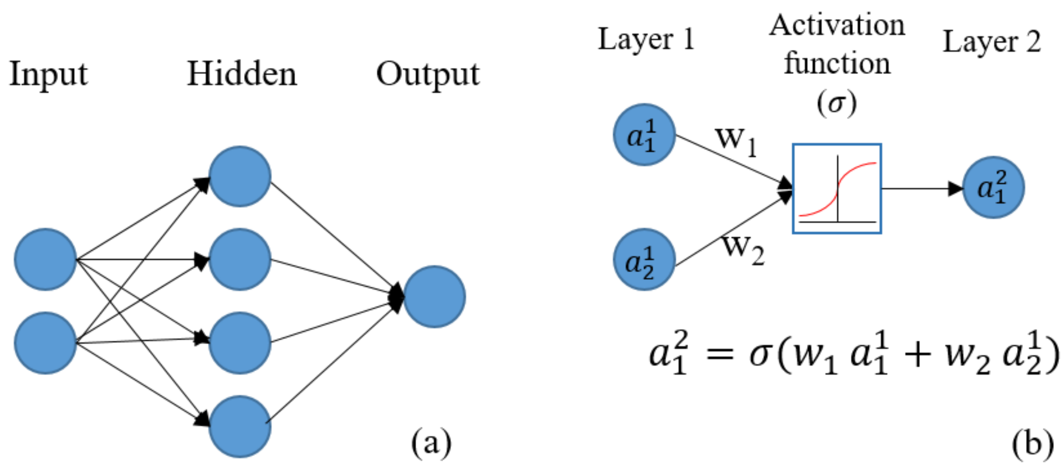


Figure 2.3: Key constituents of a CNN. (a) Architecture of a very simple CNN containing an input layer (two nodes), hidden layer (four nodes) and output layer (one node). (b) Outline of how the values of nodes in one layer are multiplied by their corresponding weight to derive the value of nodes in the next layer. The sum of the weight-multiplied node values from the previous layer is passed through the activation function, σ, to determine non-linear relationships.

During the training stage of a CNN, the weights connecting nodes are iteratively updated in two stages: feedforward and back propagation (Goodfellow et al., 2016). During the feedforward stage, input data is provided to the input nodes, with each input node corresponding to one pixel in the original image. The input layer node values are multiplied by the corresponding weights to generate the hidden layer node values (Xie and Tu 2015). The hidden layer node values are then multiplied by a separate set of weights to derive the output layer node values(s). When the feedforward stage is complete, an output is predicted, and this is compared to the corresponding observed output value contained within the original dataset (Xie and Tu 2015).

The second stage of CNN training, back propagation, starts with the calculated loss, or difference, between the predicted and observed output value (Kokkinos et al., 2015; LeCun et al., 1989). This loss is consecutively fed back through each layer of the CNN, updating the weight values, via gradient descent algorithms, until the input layer is reached (LeCun et al., 1989). The feed forward stage then commences again using the updated weights. An epoch refers to one feedforward and back propagation cycle, and CNNs are commonly trained over hundreds or thousands of epochs. CNNs are a form of neural network (NN), which all contain the abovementioned feed forward- back propagation stages. Compared with other forms of NN, CNNs contain an additional process, convolution, which is relevant to image processing (Kattenborn et al., 2021; Xie and Tu, 2015).

In convolution, a kernel is passed over an image using a pre-set function to generate a new output image (Dumoulin and Visin, 2016). The convolution kernel is placed in the top left-hand corner of the input image (Figure 2.4 (a)). The value of each pixel in the convolution kernel is multiplied by the corresponding pixel value in the input image. These multiplications are summed to derive the pixel value in the output image (Figure 2.4 (a)). This process is repeated as the kernel iteratively passes to the right (Figure 2.4 (b) – (d)), and on subsequent rows to produce the final output image (Figure 2.4 (e); Dumoulin and Visin, 2016). CNNs convolve thousands of kernels with different weights and dimensions over the input image. LeCun (2015) provides a good overview of NN, CNN, deep learning and their constituents.

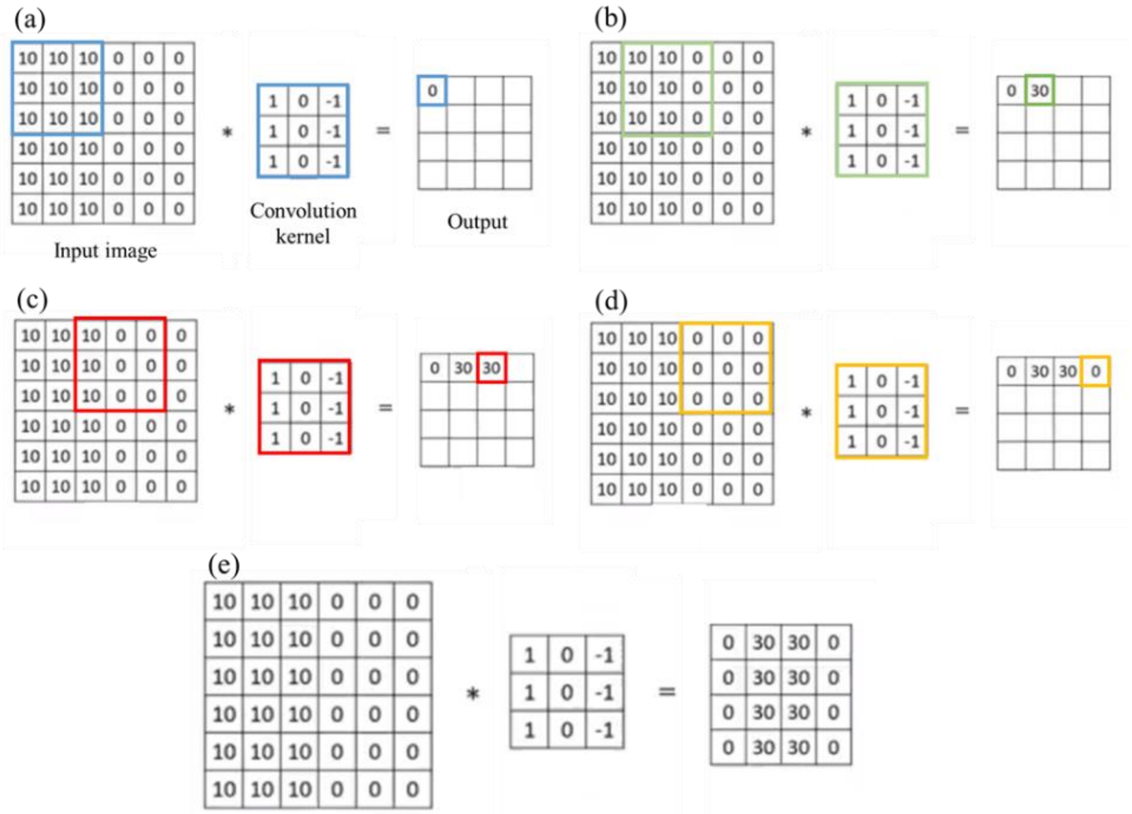


Figure 2.4: Overview of convolution. (a) – (d) kernel iteratively passes over image. (e) final output. In this example, the weight values contained within the convolution kernel result in

the higher values in the output image corresponding to edge locations in the original input image

The iterative feedforward-back propagation process, combined with convolutions, can provide CNNs with semantic information, whereby the repeated comparison between the prediction and observed output layer enables a CNN to exclusively detect features of interest, instead of identifying all features with similar spectral properties (Kokkinos et al., 2015). This feature of CNNs could be particularly important when detecting edges in remote sensing which contain a high density of edges. The onerous nature of CNN training means it is beneficial to initially determine whether SVM or non-ML techniques can detect the coastal vegetation edge.

2.6. Machine learning applications to coastal risk

The main ML tools applied to date in coastal risk studies are described in Sections 2.3 - 2.6. The complex, multifaceted nature of coastal zone dynamics (Section 1.1), combined with the recent increase in Big Data pertaining to coastal risk (Section 1.3), has prompted studies investigating whether ML tools can improve our understanding of coastline position and coastal population dynamics (Goldstein et al., 2019). This chapter individually discusses applications of ML to calculate population distributions (Section 2.7.1), and coastline position and dynamics (Section 2.7.2). Studies which have integrated multiple ML tool outputs to calculate relative coastal risk are then highlighted in Section 2.7.3. Persisting research gaps are summarised in the thesis rationale (Section 2.8).

2.6.1. Machine learning applications to coastal zone population dynamics

Fine spatio-temporal resolution datasets of population distributions in the LECZ are required to inform operational level risk management decisions (Le Cozannet et al., 2020). Census data provides population data with high veracity and exhaustivity, but census data is commonly only collected every decade, and when made publicly available, is heavily spatially aggregated. Rapid population growth and migration in the LECZ (McGranahan et al., 2007; Neumann et al., 2015), means census data can quickly become outdated. This has sparked interest in applying ML tools to alternative datasets to generate finer resolution, real-time datasets of coastal zone population distributions (Stevens et al., 2015). This section discusses three applications of ML tools to generate high resolution coastal population datasets: (i) RF to identify relationships between population vulnerability indices and the recorded impacts of storm events; (ii) natural language processing to generate real-time population distributions using social media; and (iii) dasymetric modelling, using RF algorithms, to disaggregate census data using auxiliary datasets.

The vulnerability of human populations to coastal erosion or flooding has traditionally been calculated by numerically combining multiple indices, perceived to represent vulnerability, including the age, wealth and health of the population (Cutter et al., 2003). The accuracy of output vulnerability values is, however, affected by subjectivity in vulnerability metric selection, weighting and method of combination (Willis and Fritton, 2016; Bukvic et al., 2020). More recently ML tools, including RF, have been applied to identify the statistical relationships between different vulnerability indices and the recorded economic damage (Yoon and Jeong, 2016; Heβ et al., 2017) or building damage (Foti et al., 2015) caused by historic coastal hazards. The relative importance of vulnerability indices in determining damage values has been ascertained by iteratively removing variables during ML training (Dwyer et al., 2004; Boruff et al., 2005; Abson et al., 2012; Yoon and Jeong, 2016; Uddin et al., 2019). Whilst these inductive ML-based approaches are insightful, they cannot be trained on impacts which cannot be enumerated, for example restricted access to healthcare facilities (Meyer et al., 2013). Further, these methods retrospectively determine population demographics most acutely affected by a historical coastal hazard, but do not estimate total populations at risk of future events. To estimate population distributions at risk to future coastal hazards, recent studies have used social media to provide real-time information on population dynamics.

Natural Language Processing (NLP) is a separate form of ML which uses social media data to estimate coastal population distributions and their vulnerability in real-time. NLP has been used to identify the age and gender of people affected by a coastal hazard (Mandel et al., 2012; Tellman et al., 2020), determine coastal tourist hotspots and changes in their population density between seasons (Deville et al., 2014; Li et al., 2016), and generate a time-series trajectory of the locations of people impacted by a storm as it passes over a region (Sit et al., 2019). Twitter usage has depicted flood extent, and the population impacted, in locations devoid of river gauges or other monitoring equipment (Smith et al., 2017). NLP can not, however, detect all vulnerable individuals because social media is not used universally, resulting in some flooded locations not being referred to in any Tweets (Pollard et al., 2018; Smith et al., 2017). The average social media user is better educated, wealthier, younger and more physically mobile than the general population (Li et al., 2012; Wang et al., 2019), and social media is more commonly used by tourists than local residents (Martín et al., 2020). Using NLP approaches in isolation can, therefore, result in an under-representation of the spatial impacts of coastal hazards on more vulnerable groups (Boyd and Crawford, 2012; Wang et al., 2019). The lack of exhaustivity in NLP datasets has led to research continuing to apply ML tools to disaggregate census data, which is renowned for capturing information on virtually the entire population.

Dasymetric modelling redistributes aggregated census population density data into finer resolution gridded estimates by utilising relationships between population density and other ancillary datasets (Nagle et al., 2015; Tatem et al., 2017). Traditional dasymetric models have used conventional statistical methods to apply areal weighting factors to different land cover classes, i.e. giving urban areas a high weighting and forested areas a low weighting respectively (Wu et al., 2005; Azar et al., 2010). In some settings, strong correlations between population and land use have allowed the total population to be proportionally disaggregated within an authority boundary (Wu et al., 2005; Azar et al., 2010). However, unrealistic population distributions are produced when models cannot distinguish between urban surfaces and other land covers, primarily soil and sand (Wickham et al., 2013; Nagle et al., 2014). This has prompted the use of ML tools in dasymetric modelling to disaggregate census population data using a greater diversity of ancillary datasets, including surface roughness maps, night lights, climate zone information and proximity to infrastructure and healthcare facilities (Stevens et al., 2015).

RF have been used to combine multiple spatially explicit ancillary datasets to generate the WorldPop maps, which are global population density maps at annual temporal, and 100 m^2 spatial, resolution (Sorichetta et al., 2015; Stevens et al., 2015; Gaughan et al., 2016; Sinha et al., 2019). These RF-based dasymetric models significantly outperform dasymetric models exclusively using land cover and urban extent maps (Stevens et al., 2015; Gaughan et al., 2016; Bai et al., 2019). Validating model performance, however, depends on the availability of ground-referenced municipal population data, meaning the tool predictions are potentially skewed towards generally richer countries where validation data is available (Sinha et al., 2019). Whilst these limitations require consideration, ML-derived gridded population datasets have enabled estimations of total global populations living in the LECZ (Kulp and Strauss, 2019). When compared to other datasets, WorldPop provides more accurate estimates of the number of people affected by previous coastal flood events (Reguero et al., 2019) or total populations benefitting from flood protection services provided by intertidal habitat (Menéndez et al., 2018). This highlights the potential of integrating WorldPop datasets with ML-based datasets pertaining to coastline position, to determine populations at risk to coastline change.

By considering the methods utilised, and outputs produced, by the three abovementioned techniques to produce finer resolution layers of population dynamics compared with census data, it was decided to use population layers generated via dasymetric modelling in this thesis. This was primarily due to dasymetric modelling utilising datasets that are consistently available globally, including multispectral remote-sensing imagery for land use classification and night lights (Stevens et al., 2015). In comparison, using RF to identify the relationship between vulnerability metrics and coastal hazards is dependent upon the availability of *in-situ* data pertaining to the socio-economic impacts of coastal hazards. Likewise, NLP inconsistently represents different coastal communities, with those who less commonly use social media, such as the elderly, also likely to be the most vulnerable communities (Li et al., 2012; Wang et al., 2019). An additional benefit of dasymetric modelling is that it can produce consistently gridded outputs of population dynamics (Tatem et al., 2017). This form of data is well suited to being integrated with datasets pertaining to different aspects of coastal risk, including, most appropriately for this thesis, rates of shoreline change.

2.6.2. Machine learning applications to coastline change

ML applications to coastal change have primarily focussed on (i) determining coastline position; (ii) hindcasting and predicting coastline erosion and accretion, and (iii) identifying the most important meteorological, hydrological, geomorphological or anthropogenic external forcing factors driving observed coastline change (Goldstein et al., 2019). The main non-ML tools applied to automatic coastline detection via multispectral remote sensing imagery are summarised in Section 1.4.1. This section details applications of ML tools to this task, before summarising ML tools applied to hindcasting and predicting shoreline change. Shoreline detection is an example application of image-based edge detection, a well-established area of research in computer vision (Arbelaez et al., 2007; Stanford Vision Lab, 2016). In computer vision research, edges contained within red-green-blue (RGB) images of everyday objects (e.g. house, car, cat, dog) have been identified (Simonyan and Zisserman, 2015) but additional challenges exist when applying the same task to multispectral remote sensing imagery. Remote sensing imagery contains more spectral bands, more noise and a higher density of edges compared to natural images (Liu and Jezek, 2004). The multi-dimensional nature of remote sensing imagery has sparked interest in using ML tools to automatically identify shorelines from them.

Analogous to non-ML applications, most ML-based automated shoreline detection methods have extracted the instantaneous waterline from remote sensing imagery (Toure et al., 2019). Previous studies have used RF (Bayram et al., 2017; Demir et al., 2017) and SVM (Kalkan et al., 2013; Zhang et al., 2013; Choung and Jo, 2017; Elnabwy et al., 2020) to categorise remote sensing images into land and water pixels and assign the waterline position as the boundary between the two surface cover classes. Choung and Jo (2017) found the mean error in waterline position to be lower using SVM compared to NDWI threshold contouring but SVM outputs contained a lot of ‘speckle’, attributable to the similar spectral properties of shallow water, sand and rock surfaces. The application of RF by Demir et al. (2017) identified a continuous waterline but large mean errors (greater than 22 m) were recorded between manually digitised shorelines and RF derived shorelines, attributed to noise contained within the input images. Elsewhere, SVM and RF classification performance has been adversely affected by heterogeneity in the spectral properties of water between images, caused by differences in atmospheric scattering, solar radiation incidence angle and azimuth (Chen et al., 2014). SVM and RF also rely on feature selection, including choosing the correct spectral bands to use during training, reducing the autonomy of the tools (Yu et al., 2017). These difficulties have resulted in increased attention being placed on using CNN in edge detection.

CNN techniques have also been used to automatically extract the instantaneous water line from coastal remote sensing imagery (Yu et al., 2017; Li et al., 2018; Liu et al., 2019; Erdem et al., 2021). CNN have commonly outperformed SVM and RF at edge detection and classification tasks in remote sensing imagery (Chen et al., 2014; Gao et al., 2018; Sothe et al., 2020). High CNN performance has been attributed to their use of moving kernels, meaning they simultaneously consider a neighbourhood of pixel values, rather than conducting pixel-wise classification (Gao et al., 2018). This enables CNN to detect scale-invariant features, whereby features and their edges will be in the same location, irrespective of the size of the kernel convolving over the image (Chen et al., 2014; Sothe et al., 2020). Noise and speckle are only likely to be considered as potential features when using smaller kernels and so are discarded when larger kernels convolve over the image (Sothe et al., 2020). Deeper CNN, which convolve kernels with a greater range of sizes over the image, have outperformed shallower CNN because they can detect features at different scales (Hasan et al., 2019).

While CNN provide potential in outperforming other ML tools in detecting the edges of coastal features, their reported high performance must be considered within the context of their limitations. Most pertinently, CNN require very large training datasets and are prone to overfitting when trained on small datasets (Chen et al., 2016), resulting in SVM and RF outperforming CNN when trained on small datasets (Liu et al., 2018). CNN are slower to train and more difficult to interpret than SVM and RF (Maxwell et al., 2018; Rudin, 2019). The time-consuming, black box nature of CNN training means it is logical to initially determine whether non-ML tools, or more conventional ML tools such as SVM, are capable of exclusively detecting the coastal vegetation edge. Appropriate ML tool selection is essential to accurately detect shoreline positions; these shorelines can then be used by separate ML tools to hindcast or predict future shoreline position.

ML tools have been used to hindcast and predict shoreline change at multiple spatial scales from sediment flux on sandbars (Pape et al., 2007), and beach response to storm events (Hashemi et al., 2010; Plant and Stockton, 2012; Wilson et al., 2015; Lopez et al., 2018), to national scale shoreline erosion and accretion (Giardino et al., 2019; Goldstein et al., 2019). Shorelines derived from ground-based measurements (Dickson and Perry, 2016; Wilson et al., 2019), or abovementioned ML techniques (Calkoen et al., 2021), have been used to generate a time series of antecedent shoreline positions at discrete transects. These time series have been combined with data on historical meteorological or hydrological conditions and human shoreline modification factors to hindcast, or predict, shoreline response to extraneous forcing events (Beuzen et al., 2018; Goldstein et al., 2019). At all spatial scales, ML tool selection has been shown to affect prediction performance (Dickson and Perry, 2016; Montaño et al., 2020). In one of the few studies to compare the performance of different ML tools, SVM and RF performance was identified to be less sensitive to training dataset sample size than BN and NN when predicting landslide positions (Perry and Dixon, 2018). These findings reinforce the necessity to develop a large dataset for CNN training and to compare the performance of different ML tools.

Alongside tool selection, input parameter selection has also been shown to be an important determinant of ML performance. Anthropogenic factors have commonly been highlighted as the most important factors influencing the predictive capability of ML tools (Wilson et al., 2015). Anthropogenic beach nourishment parameters most significantly influenced the capability of BN to predict future waterline position along sandy beach coastlines (Wilson et al., 2015; Wilson et al., 2019; Giardino et al., 2019). Lower BN performance in predicting rates of cliff retreat was attributed to the lack of training data available pertaining to human interventions (Hapke and Plant, 2010). Changing the parameters used to train ML tools has also provided insights into the most important parameters affecting shoreline change (Beuzen et al., 2018). BNs identified wave power and antecedent beach width to be the most important predictors of shoreline change (Beuzen et al., 2018). These principles were already well established using numerical modelling approaches (Harley et al., 2009), but BNs were able to quantify the relative importance of these factors (Beuzen et al., 2018). Parameter selection is, therefore, an important determinant of ML performance when ML tools are used to detect or predict shoreline position.

In summary, applications of ML to shoreline detection have primarily focused on the identification of the waterline. This is attributed to its consistent presence along all coastlines globally, and the close relationship between the waterline and the widely accepted definition of the coastline as the interface between the land and the sea. The relative lack of automated methods to automatically detect the coastal vegetation edge is attributed to the inconsistent presence of vegetation in the coastal zone, and heterogeneity in the coastal vegetation species and other properties, causing variability in the spectral properties of coastal vegetation (Belluco et al., 2006). Despite these challenges, there is potential for the position of the coastal vegetation edge to represent backshore dynamics not detectable using the waterline (Pollard et al., 2020), justifying further investigation into the development of automated tools in this thesis.

2.6.3. Combining multiple aspects of risk using ML

Sections 2.7.1 and 2.7.2 discuss applications of ML tools to coastal population dynamics and shoreline change in isolation but calculations of populations at risk from coastal change require the integration of datasets pertaining to natural hazards, receptor exposure and vulnerability (McLaughlin and Cooper, 2010; Bukvic et al., 2020). Multi-criteria assessments combine these dataset values (e.g. by weighted summing or multiplication) to identify 'risk hotspots', or locations at greatest risk of coastal flood and erosion (Viavattene et al., 2015; Ferreira et al., 2016; Christie et al., 2018). These studies have integrated datasets derived from statistical-based models (Hurdle and Stive, 1989; Chang et al., 2011), hydrological modelling (Jana and Hedge, 2016; Christie et al., 2018; Jäger et al., 2018; Jiménez et al., 2018), or data clustering methods (Camus et al., 2011; Villatoro et al., 2014; Calil et al., 2017; Bukvic et al., 2020). The use of thematic layers derived from non-ML tools may impact upon the scalability of the risk calculations. For example, the performance of hydrological modelling can reduce when transferred to separate coastlines, particularly when large differences in physical conditions exist between locations (Heuvelmans et al., 2004; Broderick et al., 2016; Yang et al., 2019).

In the only known study to use multiple thematic layers derived from separate ML tools, thematic layers of historic tide and rainfall data, coastal flood extent and projected urban extents were integrated to predict future risk scores of population to flooding across South Korea (Park and Lee, 2020). The ability to update values relating to all aspects of risk as new data becomes available, and to transfer the trained ML tool to other coastlines, was highlighted as a benefit of using ML tools to generate multiple thematic layers (Park and Lee, 2020). The above mentioned study, however, focussed exclusively on flood risk. The potential to integrate multiple ML-based datasets pertaining to coastline change and population dynamics, to ascertain levels of populations at risk from coastline change, has not been examined to date and is, therefore, investigated further within this thesis.

2.7. Thesis rationale

This chapter and the preceding one have detailed research relating to the extraction of the coastline from multispectral remote sensing imagery. The generation of ML and non-ML tools for the automatic extraction of the coastal waterline has been well researched (Boak and Turner, 2005; Toure et al., 2019), but automated methods to extract the coastal vegetation edge via remote sensing imagery remain largely unexplored. Despite their recorded success in detecting the instantaneous water line (Liu et al., 2019), and classifying vegetated land covers from remote sensing imagery (Maxwell et al., 2018), CNN have not been applied to the detection of coastal vegetation edges. This is attributed to the relatively understudied nature of the coastal vegetation edge, compared with other shoreline proxies such as the waterline (Toure et al., 2019). Each shoreline proxy provides representation of different processes occurring at the transition between land and sea, and overdependence upon one proxy, such as the waterline, generates an oversimplified representation of coastal processes (Pollard et al., 2020). This thesis explores the development of tools to automatically detect the coastal vegetation edge. Tool outputs are extracted from time series of imagery of different coastal locations to investigate whether the proxy can detect different coastal zone processes compared with the waterline.

Few studies compare the performance of ML and non-ML tools in detecting features in remote sensing imagery; but the most accurate, robust tool for detecting the coastal vegetation edge by other researchers and coastal stakeholders can only be determined through these comparisons. Non-ML methods (Section 1.4.1) are relatively straightforward to implement and explain, and threshold contouring in particular has been extensively applied to instantaneous waterline detection (Pekel et al., 2016; Vos et al., 2019b), but their performance at detecting the coastal vegetation edge has not been examined. One property universal in all coastal vegetation edges globally, which has never been considered to aid vegetation edge detection, is that it is the vegetation situated closest to the waterline. For this reason, this thesis creates a new non-ML based tool, which aims to identify the position of the coastal vegetation edge in remote sensing imagery by considering the proximity of the vegetation line to the waterline. The performance of this tool in detecting the coastal vegetation edge in remote sensing imagery is initially compared with other non-ML techniques, namely threshold contouring and kernel-based operators.

ML has proven successful in detecting edges and features in natural RGB images (Guo et al., 2018), but applications using remotely sensed imagery are in their infancy. SVM, RF and CNN are the main forms of ML used previously used in shoreline detection studies, with SVM highlighted as the tool least affected by training sample size (Dickson and Perry, 2016; Perry and Dickson, 2018). This attribute may be beneficial in the environmental sciences where dataset sizes are still relatively small compared with other research domains. CNN provide potential benefits over more conventional ML tools such as SVM, including their ability to detect scale-invariant features (Dumoulin and Visin, 2016; Xie and Tu, 2015), and their ability to assimilate semantic information, enabling a CNN to exclusively identify edges of interest, instead of all edges (Mountrakis et al., 2018; Liu et al., 2019). CNN have many drawbacks, however, especially the onerous training requirements and black-box nature (Maxwell et al., 2018; Rudin, 2019). It is logical, therefore, to initially investigate the performance of non-ML tools and conventional ML tools, namely SVMs, at automatically detecting the coastal vegetation edge.

Section 2.7.2 highlighted how input parameter selection can increase ML tool performance and increase transparency in how the ML tool derives its output values. This has been achieved by, for example, determining the most important parameters required by a ML tool to predict rates of shoreline change (Beuzen et al., 2018). This approach to parameter selection has not been applied to ML studies detecting shoreline position using remote sensing imagery. When using multi-spectral imagery, it is commonly not possible to train the ML tool using all available spectral bands (Kokkinos et al., 2015; Liu et al., 2019), necessitating the selection of remote sensing bands prior to ML training. Further investigation is therefore conducted in this thesis to determine the influence of different spectral band combinations on ML performance in shoreline detection tasks.

Applications of ML tools to coastline detection remain restricted to relatively short stretches of coastline and to locations which commonly already benefit from ground-referenced measurements. The longest stretch of shoreline to which ML tools have been applied is the 100 km long waterline of the heavily urbanised Jiaozhou Bay, China (Liu et al., 2019). However, most studies hindcasting and predicting shoreline change have commonly used study sites shorter than 100 km in length. The ability to apply ML tools to detect coastline position at supra-national to global scales, or to transfer trained ML tools to other locations, are widely highlighted as key benefits of ML over ground-referenced studies (Elnabwy et al., 2020) but this has yet to be fully tested. Initial trials of ML tools for smaller-scale coastlines rich in ground-referenced measurements is necessary to validate tool performance but subsequent applications to larger areas, or areas not benefitting from a wealth of ground-referenced measurements, remains limited. The ability of ML tools to detect the coastal vegetation edge at a supra-national scale is investigated in this thesis. Subsequent shoreline change analysis determines whether there are particular coastal dynamics that are only discernible when studying shoreline change at these larger spatial scales.

The combination of thematic layers derived from multiple ML tools to ascertain relative levels of risk in the coastal zone remains understudied (Section 2.7.3). This thesis identified one example study conducting this task (Park and Lee, 2020), but this study considered flood risk in isolation and did not consider risks posed by coastline change. Coastline position and morphology has been identified as an important constituent of both coastal flooding and erosion risk (Dawson et al., 2009; Möller et al., 2014) making the inclusion of coastline change data in coastal risk studies an important consideration. This thesis integrates the outputs of two separate ML tools relating to ascertain levels of populations exposed to shoreline change. These two ML tools include a CNN derived to automatically detect the coastal vegetation edge, and its change over time, and WorldPop, a third-party dataset, derived using RF, used to estimate population densities at 100 m gridded spatial resolution (Stevens et al., 2015).

Chapter 3: Vegetation edge detection using Support Vector Machines and non-machine learning approaches

3.1. Introduction

The coastline is commonly defined as the interface between the land and the sea but the complex and hierarchical nature of coastal zone processes means that they cannot be adequately represented by one single proxy (Pollard et al., 2019a). Visual proxies can include the instantaneous water line, vegetation line and beach berm line (Boak and Turner, 2005). As discussed in chapter 1 and 2, the waterline is the dominant shoreline proxy used in shoreline change studies (Toure et al., 2019) whereas methods to extract the coastal vegetation edge primarily remain restricted to manual digitisation (e.g. Leatherman, 2003; Theiler et al., 2013; McLoughlin et al., 2015). The potential for the coastal vegetation edge to represent backshore dynamics, combined with the time consuming and subjective nature of manual digitisation, justifies further investigation into the development of automated tools to detect shoreline position from multispectral remote sensing imagery.

Edge detection is a fundamental research focus in computer vision and multiple kernel-based edge detection operators have been developed that can be applied to satellite imagery, including Sobel, Roberts, Laplacian and Canny edge detection (Canny, 1986; Al-Amri et al., 2010; Toure et al., 2019). These operators identify the locations with the greatest change in greyscale intensity (Katiyar and Arun, 2014). Sobel and Roberts are examples of first derivative operators which identify locations with the steepest gradient in greyscale intensity. Second derivative operators, including Laplacian and Canny Edge detection, identify locations where the sign of the second derivative changes (Shin et al., 2001). In Canny edge detection, a thinning process, known as non-maximal suppression, subsequently removes all pixels which are not local maxima (Canny, 1986; Heene et al., 2000). When comparing operator performance, Canny edge has been shown to be more accurate than Sobel and Roberts in detecting urban edges, as a result of the operator being less sensitive to noise (Katiyar and Arun, 2014); however, comparison of edge operator performance has not been conducted on coastal scenes. Canny edge detection has been applied to waterline detection, (Heene et al., 2000; Liu and Jezek, 2004) but no investigation of edge detection operator performance to vegetation line extraction has been conducted.

More recently, supervised machine learning techniques, including Support Vector Machines (SVM) have been used to both classify coastal remote sensing imagery (Mountakis et al., 2011) and detect coastal waterline edges (Zhang et al., 2013; Chong and Jo, 2017; Elnabwy et al., 2020). There has, however, been no previous application to extracting the coastal vegetation edge. Few studies have compared the performance of different SVM kernel functions although Zhang et al., (2013) determined that polynomial SVM models outperform linear models in classifying remote sensing pixels into land and water classes. Further investigation is required to determine whether SVMs are capable of separating pixels into those both landwards and seawards of the coastal vegetation edge.

An additional method which has long been associated with identifying pixels corresponding to a vegetated land cover is the Normalised Difference Vegetation Index (NDVI). Counter-intuitively, NDVI threshold contours have been used to demarcate the coastal waterline from high resolution (less than 5 m) imagery (Dominici et al., 2019; Parente and Vallario, 2020) but the method has not been applied to the coastal vegetation edge. One key consideration with NDVI contours is the subjective choice of threshold value. An increase of 0.1 in NDVI threshold value led to a greater than 7% reduction in the area of land classified as vegetated in the Vellore District, India, for example (Gandhi et al., 2015). Further investigation is required into the influence of NDVI contour threshold value on coastal vegetated line position.

A limitation of using NDVI threshold contours, SVMs and common edge detection operators is they are unable to consider the proximity of the vegetation line to the land-water interface. There is potential that the coastal vegetation edge could be detected more accurately in high (less than 5 m) resolution imagery if consideration is given to the pattern of spectral values when traversing from the waterline inland: Most notably, the coastal vegetation line is the closest feature to the water line with elevated NDVI pixel values. This chapter thus aims to investigate and compare the performance of different tools in exclusively detecting the coastal vegetation edge. Each tool is tested on five different multispectral remote sensing images, comprising different coastal landforms. The efficacy of three separate methods is investigated: (*a*) generating NDVI threshold contours, (b) using four well-established edge detection operators: i) Laplacian, ii) Sobel, iii) Roberts and iv) Canny edge detection and (*c*) applying SVM to classify pixels landwards and seawards of the coastal vegetation edge. A novel kernel-based method, CoasTool, is then described and applied to the same images, which considers proximity to the waterline when identifying pixels corresponding to the coastal vegetation line. The performance of all methods are assessed by comparing outputs to manually digitised vegetation lines.

3.2. Methods

3.2.1. Imagery used

Images from five different locations across the UK were used to test the performance of the different edge detection tools. The sites were chosen to test the efficacy of the methods on a range of coastal settings with different shoreline curvatures: i) rocky coast headland, ii) tidal flat fronting a salt marsh, iii) rapidly retreating soft rock cliff, iv) sandy/ shingle beach dune system and v) barrier island (Table 1). This study used a combination of Planet 3 m resolution PlanetScope and 5 m resolution RapidEye multispectral imagery.

Table 3.1: Locations and key features of images used in this chapter.

Site	Dominant form of intertidal habitat	Length of shoreline in image (km)	Planet imagery used
Porthallow, Cornwall	Hard rock cliffs and bays	4.4	3 m PlanetScope
Hornsey, Essex	Salt marsh and tidal flat	4.3	3 m PlanetScope
Holderness, East Yorkshire	Soft rock cliffs with fronting sand beach	1.4	3 m PlanetScope
Dunwich, Suffolk.	Sandy shingle dune	4.1	5 m RapidEye
Blakeney Point, Norfolk	Barrier island	10.9	5 m RapidEye

3.2.2. NDVI threshold contours

The vegetation edge was also detected by contouring the NDVI of the original image at different threshold values. To investigate the sensitivity of the method to different NDVI threshold values, contours were produced at five NDVI threshold values: 0.0, 0.025, 0.05, 0.075 and 0.1. The NDVI of each image was produced using Python's *numpy* package (Equation (3.1)).

$$NDVI = \frac{NIR - R}{NIR + R} \qquad (3.1)$$

where NIR and R are the pixel values for the near infrared and red wavebands respectively.

Python's gdal.contourise() method was used to produce the NDVI threshold contours. The method initially produced multiple contour lines. To retain only the coastal vegetation line, an assumption was made that the coastal vegetation edge would be the longest contour in each image; all shorter contours and ring contours were automatically removed. Remaining output contours were subsequently imported into ArcGIS 10.5.1 and overlaid onto the original image.

3.2.3. Edge detection operators

Four pre-existing kernel-based edge detection operators were used to identify edges in the remote sensing imagery: Laplacian, Sobel, Roberts and Canny edge detection (Figure 3.1). The edge detection operators all required a single-band image as input. Two different single-band images were derived from the original multispectral imagery: i) the NDVI of the image (Equation (1)) and the greyscale image. The greyscale image was calculated as the mean of the values for red, green and blue in each pixel (Equation (3.2)).

$$Greyscale = (0.33 \times \mathrm{R}) + (0.33 \times \mathrm{G}) + (0.33 \times \mathrm{B}) \qquad (3.2)$$

Where R, G and B are the pixel values in the red, green and blue wavebands respectively.

(*a*)

-1	-1	-1
-1	8	-1
-1	-1	-1

(*b*)

-1	0	1
-2	0	2
-1	0	1

G_x

1	2	1
0	0	0
-1	-2	-1

G_y

(*c*)

1	0
0	-1

G_x

0	1
-1	0

G_y

(*d*)

2	4	5	4	2
4	9	12	9	4
5	12	15	12	5
4	9	12	9	4
2	4	5	4	2

Figure 3.1: Edge detection operator kernels: (*a*) Laplacian, (*b*) Sobel x and Sobel y, (*c*) Roberts x and Roberts y, (*d*) Canny edge detection.

The operators were applied to each image using Python's *Scipy* package. The objective of the edge detection operators is to find and elevate the value of pixels corresponding to edges, which are characterised as locations with steep gradients in pixel values (Katiyar and Arun, 2014). For the Sobel, Laplacian and Roberts operators, the output images were therefore filtered and only the pixels with the top 5% of values were retained. The Canny edge detector produced a binary raster layer with edge locations = 1 and non-edge = 0, so the Canny Edge outputs were not filtered for the top 5% of pixels. The initial output images and filtered images were subsequently imported into ArcGIS 10.5.1.

3.2.4. Validating kernel operator performance

The position of the filtered kernel operator output pixels were compared to the manually digitised vegetation line. The pixel-based evaluation metrics used were user accuracy (Equation (3.3)), producer accuracy (Equation (3.4)) and F_1 (Equation (3.5)):

$$U_{\mathrm{A}} = \frac{P_{\mathrm{True}}}{P_{\mathrm{True}}+P_{\mathrm{False}}} \quad (3.3)$$

$$P_{\mathrm{A}} = \frac{P_{\mathrm{True}}}{P_{\mathrm{True}}+N_{\mathrm{False}}} \quad (3.4)$$

$$F_1 = \frac{P_{\mathrm{A}} \times U_{\mathrm{A}}}{P_{\mathrm{A}}+U_{\mathrm{A}}} \quad (3.5)$$

where U_{A} and $P_{\mathrm{A}}$are the user accuracy and producer accuracy values respectively. P_{True} = True Positive and N_{True} = True Negative, each corresponding to correctly classified pixels and P_{False} = False Positives and N_{False} = False Negative, each corresponding to incorrectly classified pixels. U_{A} is related to the 'precision' of the kernel operator output; the value of U_{A} decreases if pixels not corresponding to the manually digitised line are identified as the vegetation edge. P_{A} represents the proportion of manually digitised pixels which were correctly identified by the kernel-operators; the P_{A} value reduces if pixels corresponding to the manually digitised line are missed.

A pixel incorrectly identified to be the vegetation line by the kernel operator will be classified as a false positive pixel, irrespective of the distance from the manually digitised line. To account for 'near-misses', where pixels elevated by the kernel operators are close to the manually digitised line, the manually digitised lines were buffered to be three (9 – 15 m) and five (15 – 25 m) pixels wide (instead of one). Relaxed U_A, P_A and F_1 scores were calculated by comparing kernel operator filtered outputs to the buffered manually digitised vegetation lines.

3.2.5. Support Vector Machines

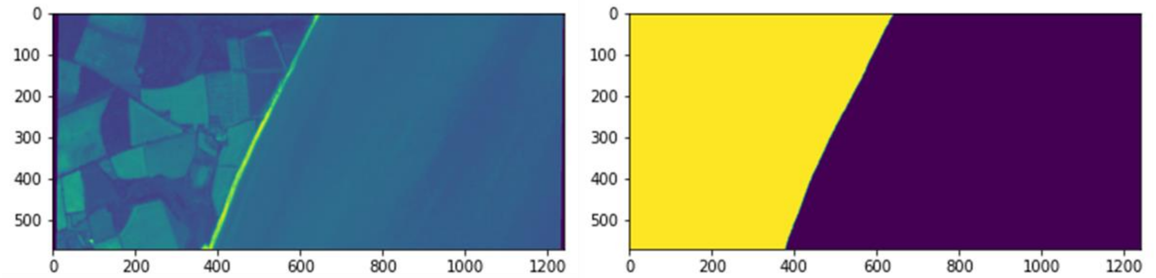
To train the Support Vector Machines (SVM) models, a training dataset was first generated which assigned all pixels as being either landwards or seawards of the coastal vegetation edge. To produce the training dataset, field-based measurements of the coastal vegetation edge were collected using a real-time kinematic global positioning system (RTK-GPS) with horizontal positional accuracy of 30 mm from three locations along the Suffolk coastline: Walberswick, Dunwich and Covehithe. RTK-GPS measurements were collected on 7 September 2019. To ensure there was a least one RTK-GPS measurement per Planet image pixel, measurements were collected every two metres, or when there was a notable change in the direction of the vegetation edge. The ground-referenced line was overlaid on 3 m PlanetScope images collected on 12 September 2019 (Figure 3.2 (*a*)). Between 7 and 12 September 2019 waves approached from a dominant north easterly direction and rarely exceeded 1 m significant wave height (maximum peak significant wave height at Southwold Approach was 1.45 m (Cefas, 2020)). Due to these wave conditions, there is a high degree of confidence that the vegetation line remained stable over this time period.

The RTK-GPS measurements were converted to raster format (Figure 3.2 (*b*)) and all pixels landwards and seawards of the vegetation line were assigned a value of one and zero respectively (Figure 3.2 (*c*)). The arrays of ones and zeros were concatenated with the original four-band planet training images. The images were subsequently unravelled so that each pixel was represented as a single five column row in a table (Figure 3.2 (*d*)). The first four columns represented the pixel values in the red, green, blue (RGB) and nir wavebands, and the fifth column stated whether the pixel was landwards (one) or seawards (zero) of the vegetation line.

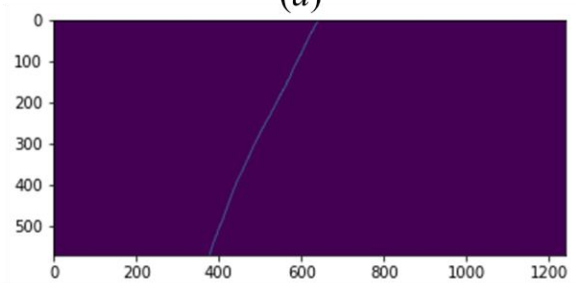
By applying the same method to all three sites containing ground-referenced measurements, over 1.1 million pixels were generated for the training dataset. Prior to SVM model training, any rows in the training table (Figure 3.2 (*d*)) corresponding to pixels with NaN values were removed using python’s *numpy* library.

SVM training and predicting was conducted using python’s *scikit-learn* library (Pedregosa et al., 2011). In light of previous findings that polynomial outperform linear SVM models (Zhang et al., 2013), the performances of a linear and polynomial SVM model were compared. For both SVM model types, performance was analysed using a regularisation value of one, two, three and four. All SVM models were separately trained using the abovementioned process and used to predict the location of pixels landwards (pixel value = one) and seawards (pixel value = 0) of the coastal vegetation edge.

To compare the position of the SVM predicted vegetation line and the manually digitised vegetation line, the SVM outputs were contoured using a value of one. The resulting contour map had to be semi-manually reviewed to remove contours not corresponding to the coastal vegetation edge, for example removing ring contours or other contours situated inland of the coastline. The validation method using the Digital Shoreline Analysis System (DSAS) is outlined in detail in section 2.6.



(*a*) (*c*)



Red	Green	Blue	Nir	0/1
4543	3709	2461	7530	1
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
6127	5341	4270	1365	0

(*b*) (*d*)

Figure 3.2: SVM training stages. (*a*) Original four-band Planet image. (*b*) Rasterised vegetation line derived from ground-referenced RTK-GPS measurements. (*c*) All pixels landwards and seawards of the vegetation line were assigned a value of 1 (yellow) and 0 (blue) respectively. (*d*) Tabulation of pixel values for SVM training. Each row represents a

different pixel in the training dataset. The first four columns provide the image intensity values for every spectral band. The fifth column specifies whether the pixel was located landwards (1) or seawards (0) of the coastal vegetation edge.

3.2.6. CoasTool

3.2.6.1.Extracting the waterline

CoasTool considers the proximity of the vegetation line to the instantaneous waterline. The first step was, therefore, to detect the approximate location of the land-water interface using a Normalised Difference Water Index (NDWI) threshold contouring method, as applied in other studies (Equation (3.6); Kuleli et al., 2011; Vos et al., 2019b). The waterline was extracted by first calculating the NDWI of each image (McFeeters, 1996). A double-peaked NDWI histogram was formed, where water pixels had a positive NDWI value close to one and land pixels had a near-zero or negative value. The image-specific threshold separating land and water pixels was determined as the NDWI intensity value with the minimum value between the two lobes. This intensity value represents the relatively low number of pixels at the transition between land and water (Vos et al., 2019b; Figure 3.3).

$$NDWI = \frac{G - NIR}{G + NIR} \qquad (3.6)$$

where G and NIR are the pixel values for the green and near infrared wavebands respectively.

The NDWI image was contoured using the image-specific threshold value and Python’s gdal.contourise() method to produce a vector polyline of the waterline. The ocean constituted the largest cluster of connected pixels above the NDWI threshold in every image. To remove waterlines surrounding inland waterbodies or clouds with high NDWI values, contours produced around smaller clusters of pixels above the NDWI threshold were automatically deleted. The pixels corresponding to the waterline were defined as those pixels which were overlaid by the remaining NDWI contour polyline (Figure 3.3).

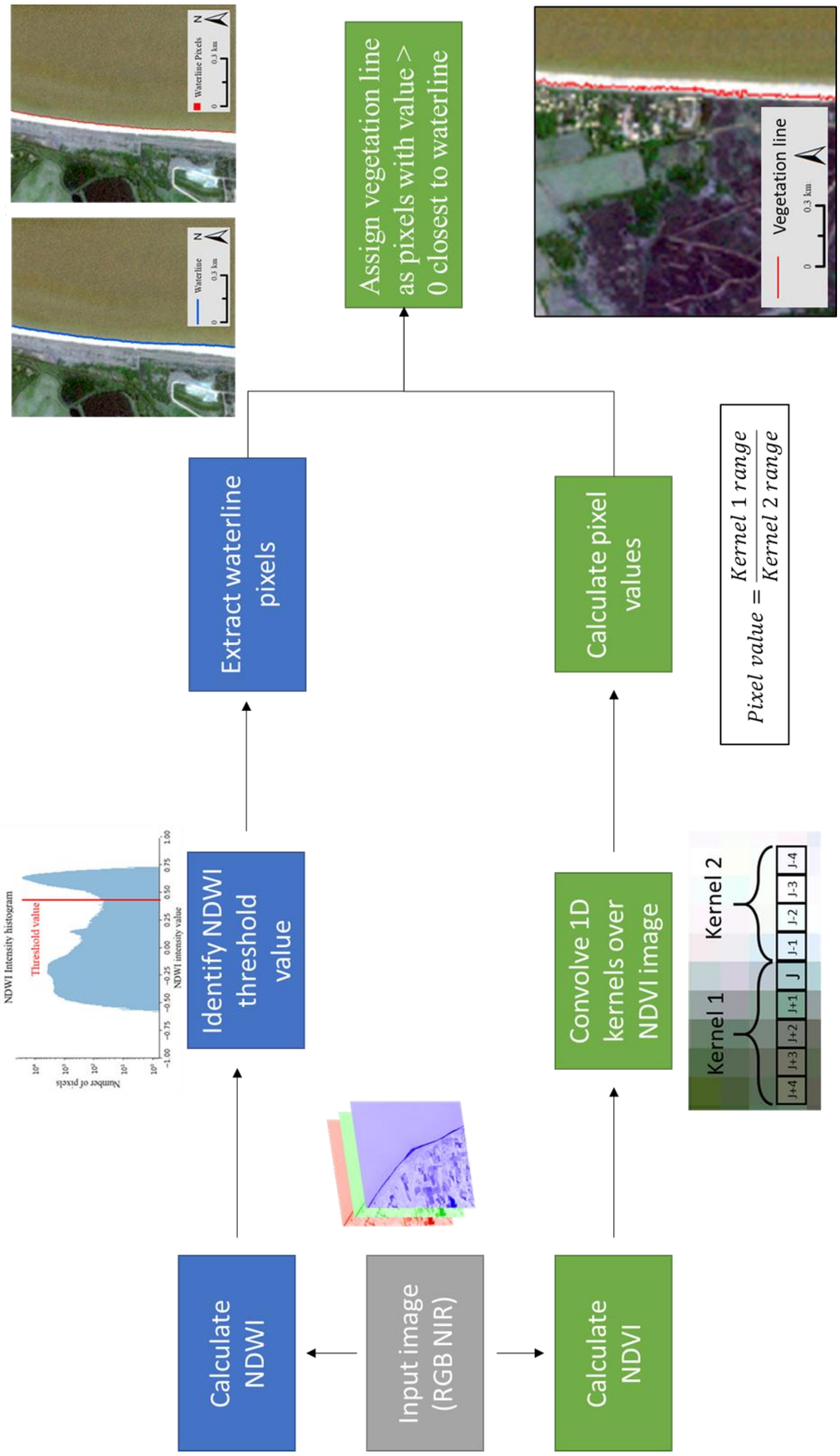


Figure 3.3: Overview of Coastal Methodology. Top: NDWI threshold method for extracting the waterline. Bottom: Method to convolve kernels over NDVI of the input image to elevate the value of pixels at the vegetation line. Right: The vegetation line is defined as those pixels with elevated values closest to the waterline.

3.2.6.2.Identifying vegetation line pixels

Prior to performing the following steps, every image was rotated so that the coastline ran approximately top to bottom, with the land on the left-hand side of the image. Two 1-D kernels convolved from left to right and top to bottom over the image with a pixel stride length of one. Kernel 1 contained the pixel of interest whose value was being calculated, j, and the four pixels immediately to the left (landwards) [j, j+1, j+2, j+3, j+4]. Kernel 2 contained the four pixels immediately to the right (seawards) of j and the value 0.0: [j-1, j-2, j-3, j-4, 0.0] (Figure 3.4). As pixels containing vegetated land covers have high NDVI values, it was hypothesised that at the vegetation line there would be a large difference between the maximum and minimum value in kernel 1, and a small difference in pixel values in kernel 2. To amplify the values of pixels at the vegetation line, the value of every pixel in the image was calculated using Equation (3.7):

$$Pixel\ value = \frac{Kernel\ 1\ range}{Kernel\ 2\ range} \tag{3.7}$$

Non-maximal suppression was then used to remove most pixels. A 1-D sliding window five pixels wide was passed over the output image. Pixels retained their value if they had the highest value in this sliding window with all other pixels being assigned a value of 0.0. To further discard pixels unlikely to be the vegetation line, all pixels within a five pixel buffer of the water line pixels were set to 0.0 and all pixels outside of a five pixel buffer of the 0.0 NDVI contour were set to 0.0.

Finally, to find the vegetation line pixels, the Euclidean distance function was used to find the nearest pixel with value greater than 0.0 from every waterline pixel. To generate a vectorised representation of the vegetation line, the resulting layer from every image was imported into ArcGIS 10.5.1. The *Raster to Point* and then *Point to Line* functions were used to convert the remaining vegetation line pixels into a vector vegetation polyline.

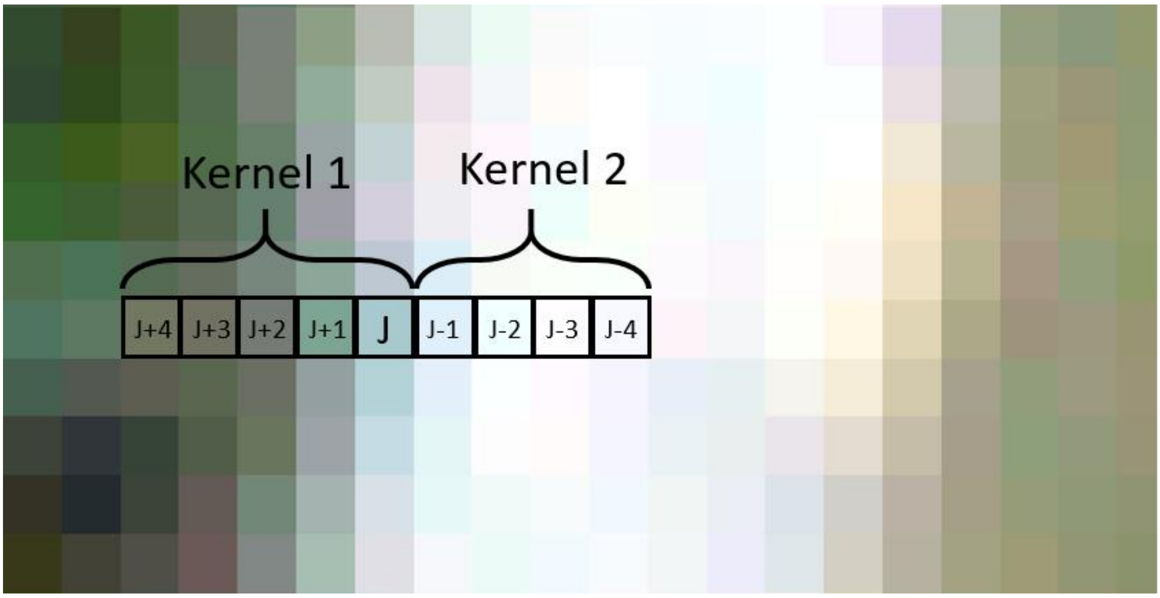


Figure 3.4: Visual representation of the two kernels used in CoasTool. Kernel 1 contained the values within the pixel of interest and the four landward pixels. Kernel 2 contained the values of the four seaward pixels. Image background shows a zoomed in section of Dunwich, Suffolk. Green pixels to the left are vegetation, white pixels in the centre are the sandy/ shingle beach and dune system and the brown pixels to the right are water.

3.2.6.3.Validation

To validate the accuracy of CoasTool, NDVI threshold contours and the contoured SVM outputs, outputs were compared to manually digitised vegetation lines. The vegetation lines in the five images were manually digitised in ArcGIS 10.5.1. To improve the positional accuracy of the manually digitised lines, visual reference was made to high spatial resolution vertical aerial imagery (50 cm) during digitisation to aid with the identification of vegetated features not readily discernible in the 3 – 5 m resolution Planet imagery. The NDVI of each image was also overlaid at high transparency to aid identification of the coastal vegetation edge.

The ArcMap plugin Digital Shoreline Analysis System (DSAS; (Thieler et al.,, 2009; USGS, 2018)) v5.0 was used in ArcGIS 10.5.1 to calculate distance between i) manually digitised shorelines and ii) CoasTool outputs and iii) NDVI threshold contours. Distances between lines were calculated along transects running perpendicular to the dominant shoreline direction at 10 m alongshore spacing. To reduce transect crossing on sinuous coastlines, each transect was drawn orthogonal to a smoothed baseline that was drawn seawards of the shoreline position. This was generated by calculating mean baseline angle over a 200 m interval, with the transect location at the midpoint.

The average error or distance between the manually digitised line and edge detection outputs across all transects was calculated using Root Mean Squared Error (RMSE, Equation (3.8)) and Mean Absolute Error (MAE, Equation (3.9)):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(o_i - p_i)^2} \qquad (3.8)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|o_i - p_i| \qquad (3.9)$$

where o_i and p_i refer to the observed and predicted vegetation line position respectively and n is the number of transects where the different between the observed and predicted lines were measured.

3.3. Results

3.3.1. NDVI threshold and threshold based methods

The NDVI contours produced using every threshold value, 0.0, 0.025, 0.05, 0.075 and 0.1, were overlaid on the original Planet image used for that site (Figure 3.5). At every site, the difference or error was calculated between each NDVI threshold contour and the manually digitised vegetation line (Table 2). NDVI threshold contours produced the smallest RMSE and MAE values at Dunwich and Holderness using the 0.0 threshold (Dunwich: MAE = 4.23 m; Holderness: MAE = 2.13 m; Table 2). The MAE values produced by the 0.0 NDVI contour were far larger at the other three sites (Blakeney Point = 107.61 m, Hornsey = 25.34 m and Porthallow = 58.02 m). The difference in error value when using different thresholds was most notable at Blakeney (0.0 contour RMSE = 186.21 m; 0.1 contour RMSE = 368.83 m). With the exception of Porthallow where the 0.0 NDVI contour produced the largest errors, the 0.0 NDVI threshold contour produced the smallest error values at every site (Porthallow 0.0 contour RMSE = 116.51 m and MAE = 58.02 m).

Table 3.2: RMSE and MAE between NDVI contours and manually digitised vegetation lines. Error values were calculated for NDVI contours using threshold values 0, 0.025, 0.05, 0.075 and 0.1.

	Blackeney Point		Dunwich		Holderness		Hornsey		Porthallow	
NDVI Value	RMSE (m)	MAE (m)	RMSE (m)	MAE (m)	RMSE (m)	MAE (m)	RMSE (m)	MAE (m)	RMSE (m)	MAE (m)
0	186.21	107.61	5.78	4.23	2.31	2.13	92.76	25.34	116.51	58.02
0.025	213.34	135.00	5.84	4.87	1.62	0.99	95.90	28.13	113.10	54.35
0.05	238.42	161.40	6.47	5.13	3.63	1.57	108.16	38.43	110.50	50.33
0.075	291.97	200.17	6.74	5.88	29.11	12.55	238.86	137.54	109.16	48.12
0.1	368.83	262.09	7.14	6.03	41.21	18.58	292.53	183.16	108.43	46.22

RMSE values were consistently larger than MAE values at every site. The largest discrepancies in RMSE and MAE values were generated by the 0.1 contours. For example at Holderness, RMSE was 0.18 m larger than MAE for the 0.0 NDVI threshold contour (RMSE = 2.31 m, MAE = 2.13 m), but was 22.63 m larger for the 0.1 contour (RMSE = 42.21 m, MAE = 18.58 m).

The insets in Figure 3.5 emphasise locations where the use of different NDVI threshold values led to large discrepancies in contour position. In general, the contours pertaining to the higher NDVI threshold values were situated further inland than the lower-valued threshold contours. At Porthallow, the 0.0 NDVI contour was situated seawards of the waterline along some transects (Figure 3.5 (*a*)). Large differences in NDVI contour position were generated for much of the marsh area at Hornsey (Figure 3.6 (*b*)), and to the north west of the image at Holderness (Figure 3.5 (*c*)). At Blakeney, it was only possible to detect the vegetation along the westerly edge of the barrier island using the 0.0 NDVI threshold contour (Figure 3.5 (*e*)).

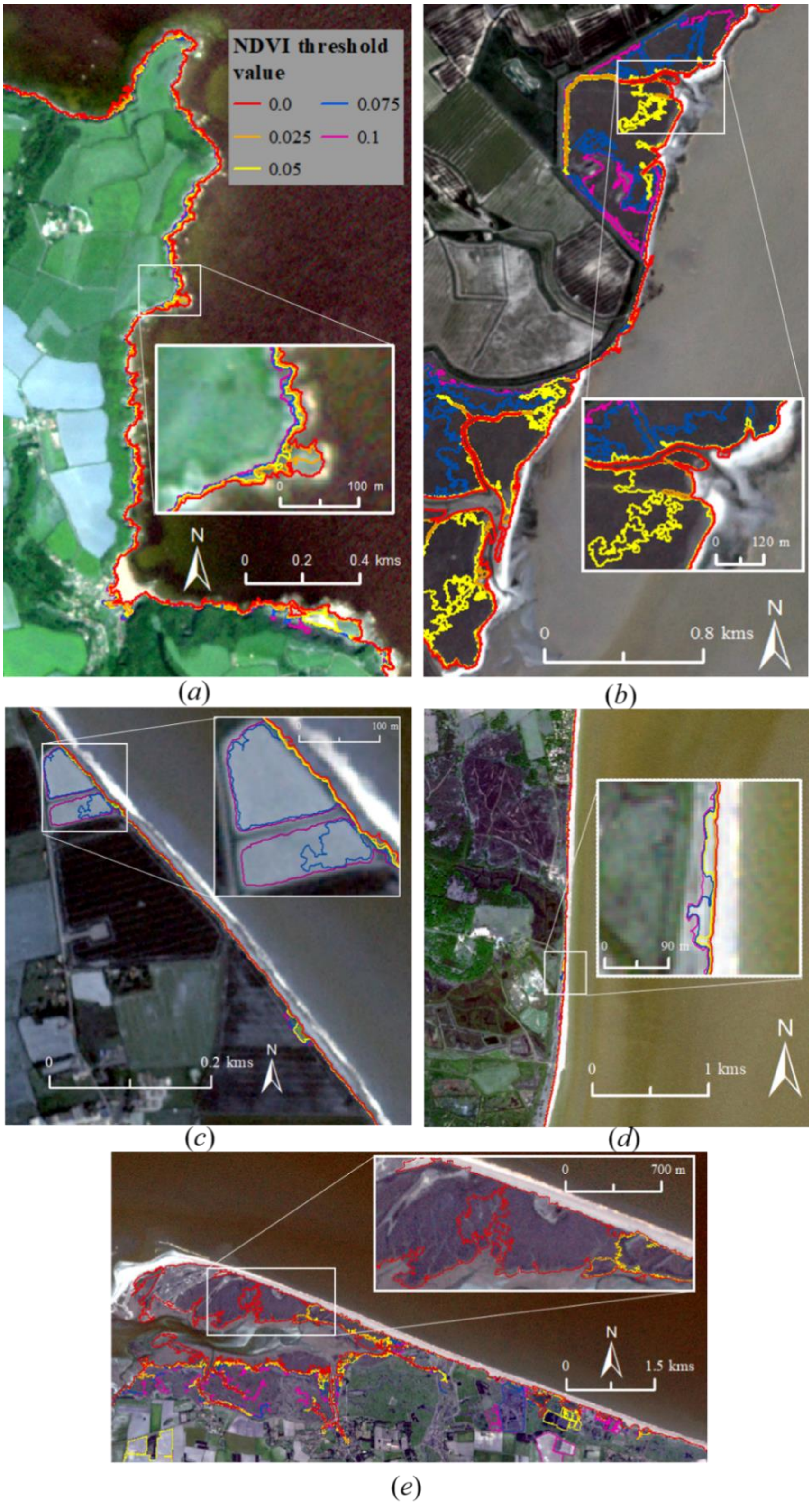


Figure 3.5: NDVI threshold contour outputs at (*a*) Porthallow, (*b*) Hornsey, (*c*) Holderness, (*d*) Dunwich and (*e*) Blakeney Point. All five images use the same colour ramp for the different contour threshold values. Insets emphasise locations with large discrepancies in contour location.

3.3.2. Edge detection operators

Edges in the NDVI and greyscale images were detected using the edge detection kernels at all five sites (Figure 3.6, 3.7, 3.8, &3.9). The edge detection operators did not perform consistently better or worse when using either the greyscale or NDVI image as input. At Holderness (Figure 3.6) and Porthallow (Figure 3.7), higher U_A, P_A and F_1 values were achieved when detecting edges in the NDVI images (Table 3). Conversely, at Blakeney (Figure 3.8), Dunwich (Figure 3.9) and Hornsey, higher U_A, P_A and F_1 values were achieved when the operators detected edges in the greyscale image (Table 3.3). See Supplemental Materials A to view all filtered and unfiltered outputs produced by the edge detection operators at the five sites when using the NDVI and greyscale images.

The greatest U_A score was generated by the Sobel and Roberts kernels over the greyscale image at Hornsey (U_A = 0.96, P_A = 0.14), meaning the kernel operators correctly identified 96% of the pixels corresponding to the manually digitised vegetation edge. The Sobel kernel was also the best performing kernel over the NDVI image at Porthallow (U_A = 0.86, P_A = 0.31). The highest P_A scores were produced by the Canny edge detector at three of the five sites (Table 3.3). For the other three kernel operators, the highest P_A score recorded was 0.31, meaning less than one out of three pixels were correctly detected as the vegetation edge. The Laplacian kernel consistently performed the worst at every site when applied to both the NDVI and greyscale image. When the vegetation line was not buffered, the maximum P_A value recorded at all sites using any kernel was just 0.13 (Canny edge detection at Holderness). For the other three edge detection operators, the largest P_A value recorded was 0.06 (Sobel edge detection using NDVI image at Porthallow).

Table 3.3: Sobel, Laplacian, Roberts and Canny edge detection operator performance at the five test sites. Operator performance when detecting edges in the NDVI and greyscale images is compared. Green boxes highlight which kernel operator performed the best at each site, and whether the operator was detecting edges in the NDVI or greyscale image. Column headers ‘1’, ‘3’ and ‘5’ correspond to the buffered pixel-width of the manually digitised vegetation line.

			P_A			U_A			F_1		
		Manually digitised line pixel width	1	3	5	1	3	5	1	3	5
Porthallow	Sobel	Grayscale	0.03	0.10	0.13	0.48	0.39	0.29	0.03	0.08	0.09
		NDVI	0.06	0.22	0.31	0.86	0.81	0.67	0.05	0.17	0.21
	Laplacian	Grayscale	0.01	0.04	0.09	0.13	0.18	0.20	0.01	0.03	0.06
		NDVI	0.01	0.02	0.03	0.09	0.09	0.07	0.01	0.02	0.03
	Roberts	Grayscale	0.03	0.11	0.15	0.49	0.43	0.35	0.03	0.09	0.11
		NDVI	0.05	0.19	0.26	0.81	0.72	0.57	0.05	0.15	0.18
	Canny	Grayscale	0.00	0.01	0.02	0.03	0.04	0.05	0.00	0.01	0.02
		NDVI	0.13	0.45	0.62	0.17	0.16	0.13	0.07	0.11	0.11
Hornsey	Sobel	Grayscale	0.02	0.09	0.14	0.96	0.9	0.74	0.03	0.08	0.11
		NDVI	0.01	0.05	0.07	0.59	0.51	0.37	0.02	0.05	0.06
	Laplacian	Grayscale	0.00	0.02	0.05	0.11	0.21	0.27	0.00	0.02	0.04
		NDVI	0.00	0.02	0.03	0.14	0.16	0.15	0.00	0.02	0.02
	Roberts	Grayscale	0.03	0.09	0.14	0.96	0.9	0.75	0.02	0.09	0.12
		NDVI	0.01	0.05	0.06	0.57	0.48	0.36	0.02	0.04	0.05
	Canny	Grayscale	0.00	0.01	0.01	0.07	0.06	0.07	0.00	0.01	0.01
		NDVI	0.03	0.09	0.10	0.26	0.18	0.12	0.03	0.06	0.05
Holderness	Sobel	Grayscale	0.01	0.06	0.10	0.25	0.25	0.23	0.01	0.05	0.07
		NDVI	0.04	0.17	0.21	0.78	0.64	0.45	0.04	0.13	0.14
	Laplacian	Grayscale	0.01	0.03	0.07	0.12	0.12	0.14	0.01	0.02	0.05
		NDVI	0.00	0.03	0.08	0.07	0.13	0.17	0.00	0.02	0.05
	Roberts	Grayscale	0.01	0.07	0.11	0.21	0.25	0.24	0.01	0.05	0.08
		NDVI	0.04	0.15	0.19	0.71	0.58	0.42	0.04	0.12	0.13
	Canny	Grayscale	0.02	0.08	0.10	0.25	0.24	0.18	0.01	0.06	0.07
		NDVI	0.14	0.46	0.67	0.20	0.15	0.12	0.08	0.11	0.10
Dunwich	Sobel	Grayscale	0.01	0.04	0.07	0.40	0.44	0.47	0.01	0.04	0.06
		NDVI	0.01	0.02	0.04	0.25	0.26	0.28	0.00	0.02	0.03
	Laplacian	Grayscale	0.00	0.01	0.02	0.12	0.11	0.11	0.00	0.01	0.01
		NDVI	0.00	0.01	0.02	0.07	0.11	0.09	0.00	0.01	0.01
	Roberts	Grayscale	0.01	0.04	0.08	0.44	0.49	0.52	0.02	0.04	0.07
		NDVI	0.00	0.02	0.04	0.25	0.26	0.29	0.01	0.02	0.04
	Canny	Grayscale	0.00	0.01	0.02	0.04	0.09	0.12	0.00	0.01	0.02
		NDVI	0.00	0.01	0.05	0.00	0.00	0.01	0.00	0.00	0.01
Cromer	Sobel	Grayscale	0.01	0.04	0.07	0.82	0.75	0.64	0.01	0.04	0.06
		NDVI	0.00	0.01	0.02	0.21	0.21	0.16	0.00	0.01	0.02
	Laplacian	Grayscale	0.00	0.01	0.02	0.19	0.19	0.18	0.00	0.01	0.02
		NDVI	0.00	0.00	0.00	0.02	0.02	0.02	0.00	0.00	0.01
	Roberts	Grayscale	0.01	0.04	0.06	0.76	0.71	0.61	0.01	0.04	0.06
		NDVI	0.00	0.01	0.01	0.16	0.13	0.10	0.00	0.01	0.01
	Canny	Grayscale	0.00	0.00	0.01	0.13	0.14	0.12	0.00	0.01	0.01
		NDVI	0.01	0.03	0.04	0.06	0.04	0.03	0.01	0.01	0.02

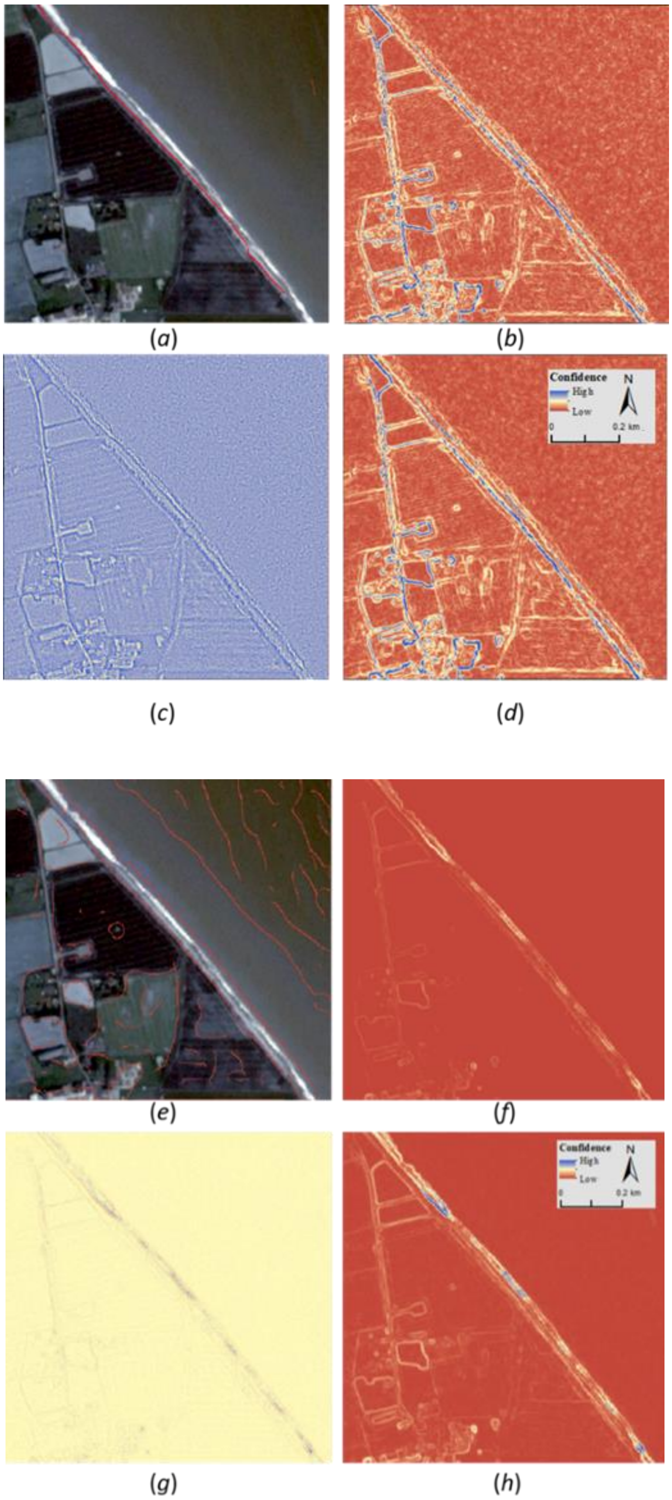


Figure 3.6: Outputs from applying edge detection operators to NDVI and greyscale image at Holderness, East Yorkshire. Outputs from applying (*a*) Canny edge, (*b*) Roberts, (*c*) Laplacian and (*d*) Sobel edge detection to the NDVI image.Outputs from applying (*e*) Canny edge, (*f*) Roberts, (*g*) Laplacian and (*h*) Sobel edge detection to the greyscale

image.Blue and red correspond to pixels determined to be an edge with high and low confidence respectively.

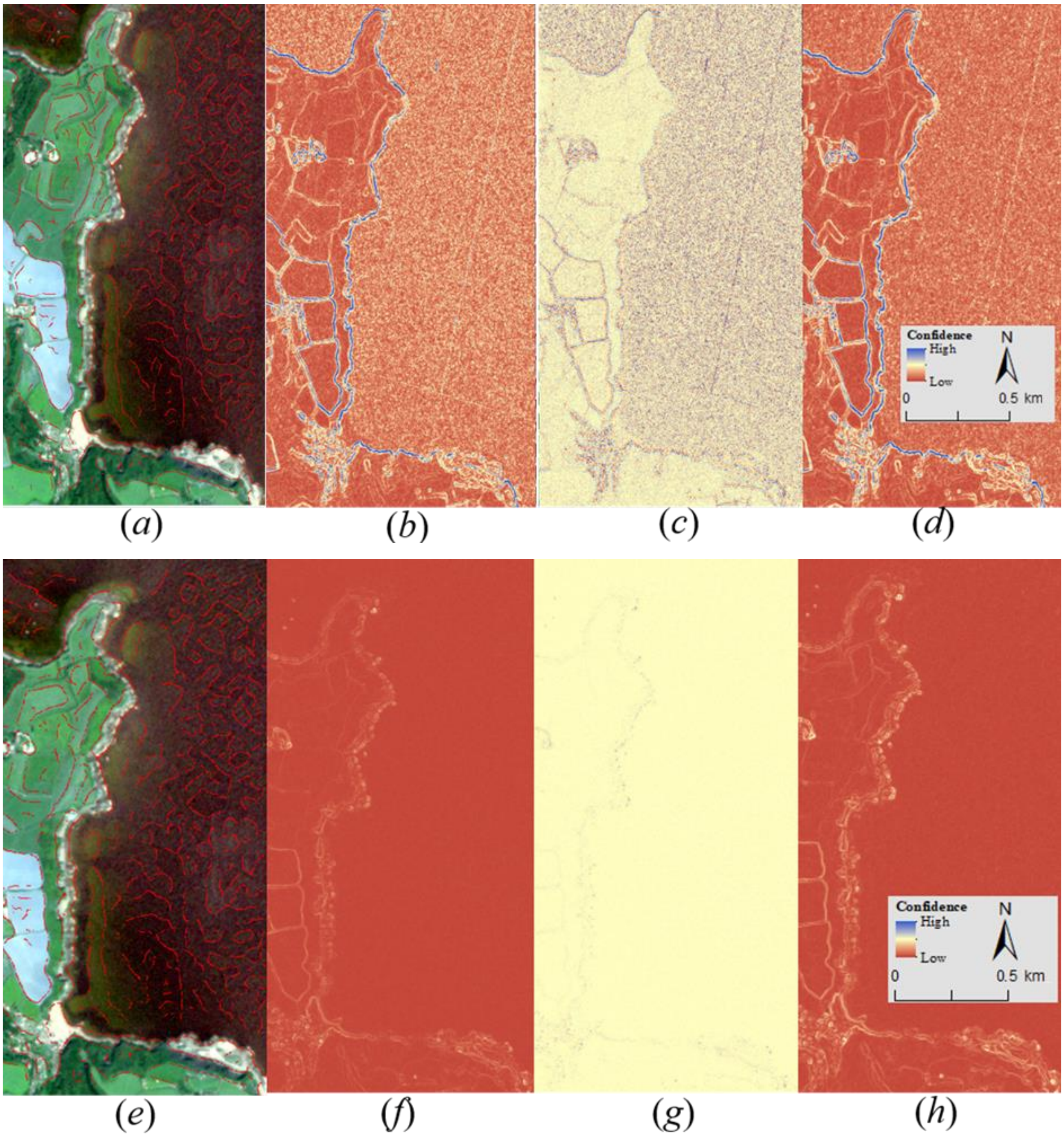
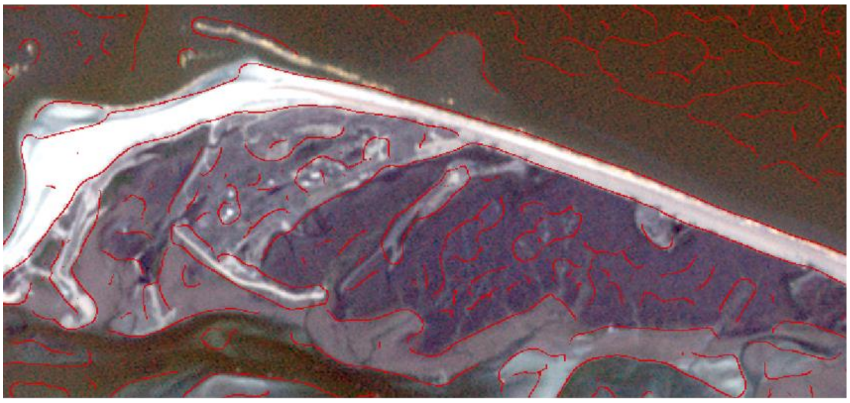
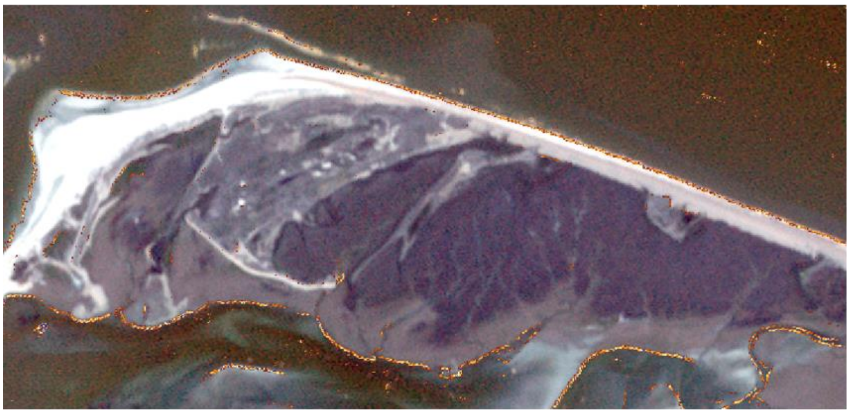


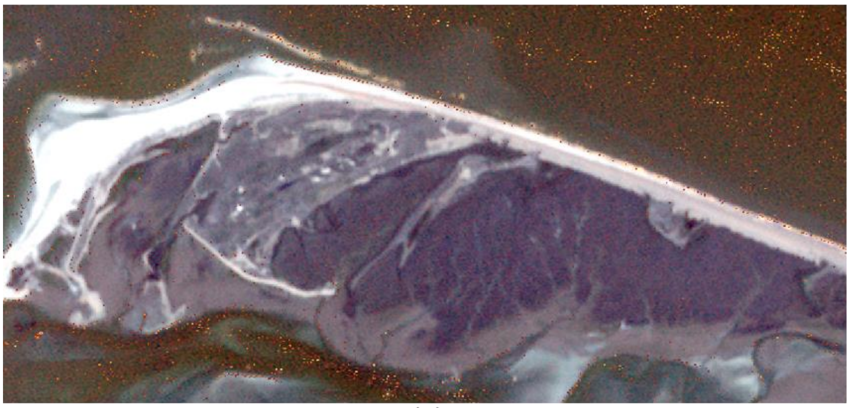
Figure 3.7: Outputs from applying edge detection operators to NDVI and greyscale image at Porthallow, Cornwall. Outputs from applying (*a*) Canny edge, (*b*) Roberts, (*c*) Laplacian and (*d*) Sobel edge detection to the NDVI image. Outputs from applying (*e*) Canny edge, (*f*) Roberts, (*g*) Laplacian and (*h*) Sobel edge detection to the greyscale image. Blue and red correspond to pixels determined to be an edge with high and low confidence respectively.



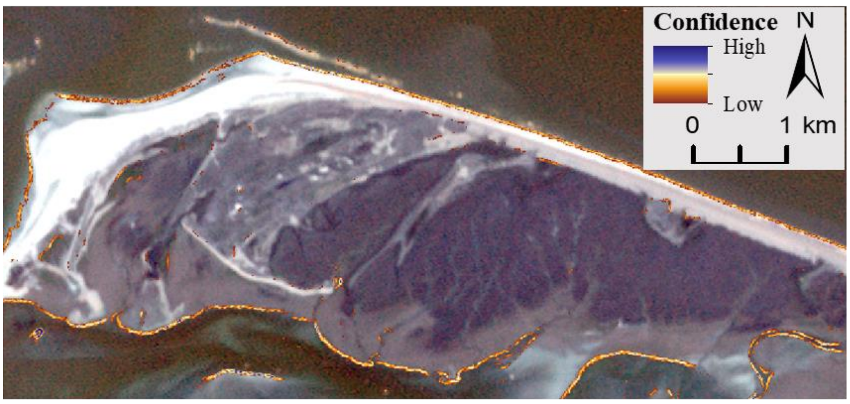
(*a*)



(*b*)



(*c*)



(*d*)

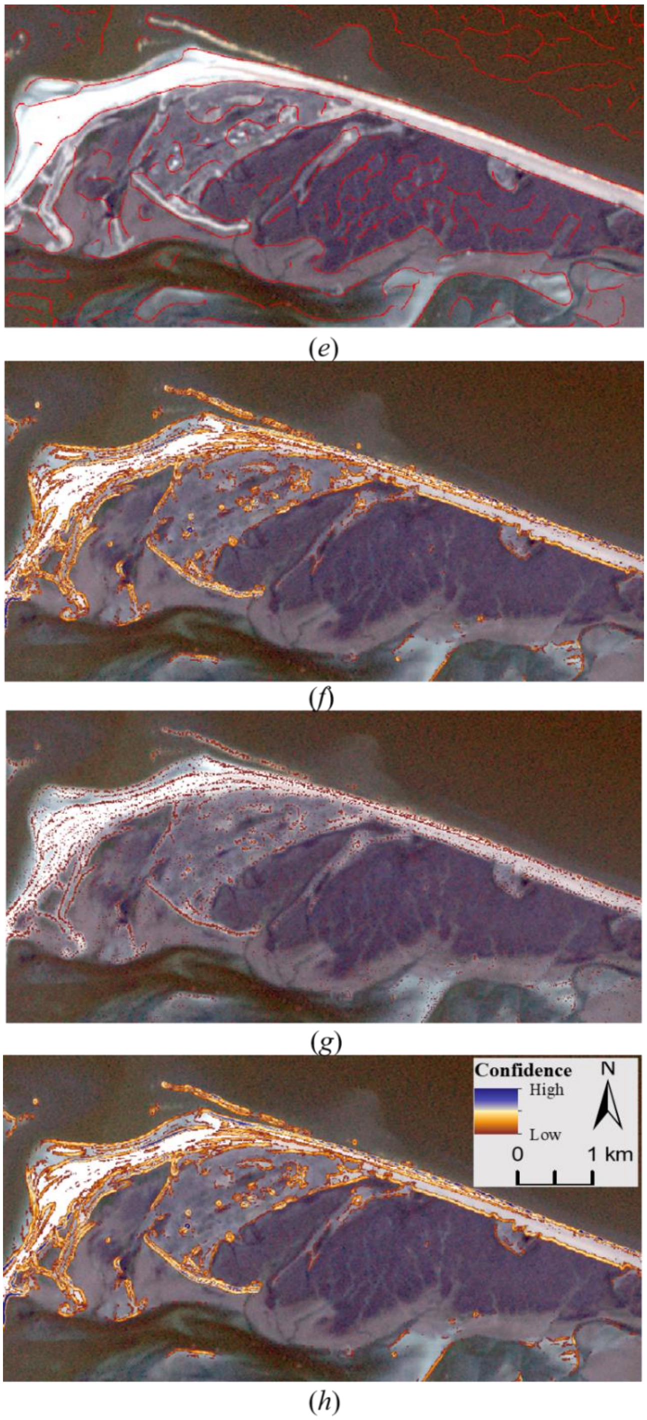


Figure 3.8: Outputs from applying edge detection operators to NDVI and greyscale image at Blakeney, Norfolk. Output images have been filtered so only the 5% of pixels with the highest values after applying the kernels are shown. Outputs from applying (*a*) Canny edge, (*b*) Roberts, (*c*) Laplacian and (*d*) Sobel edge detection to the NDVI image. Outputs from applying (*e*) Canny edge, (*f*) Roberts, (*g*) Laplacian and (*h*) Sobel edge detection to the greyscale image. Blue and red correspond to pixels determined to be an edge with high and low confidence respectively.

(*a*)

(*b*)

(*c*)

(*d*)

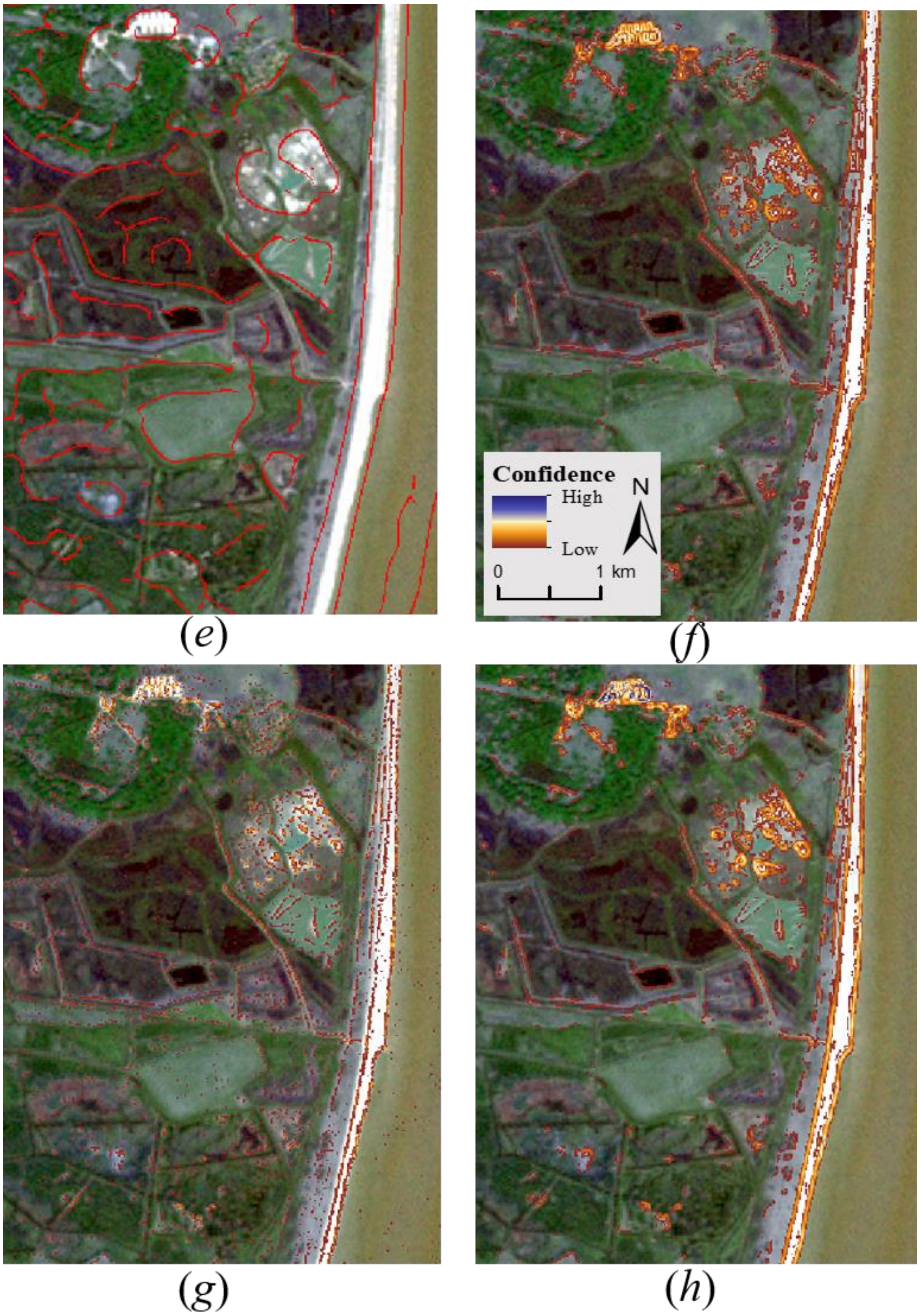


Figure 3.9: Outputs from applying edge detection operators to NDVI and greyscale image at Dunwich, Suffolk. Output images have been filtered so only the 5% of pixels with the highest values after applying the kernels are shown. Outputs from applying (*a*) Canny edge, (*b*) Roberts, (*c*) Laplacian and (*d*) Sobel edge detection to the NDVI image. Outputs from applying (*e*) Canny edge, (*f*) Roberts, (*g*) Laplacian and (*h*) Sobel edge detection to

the greyscale image. Blue and red correspond to pixels determined to be an edge with high and low confidence respectively.

3.3.3. Support Vector Machines

Linear and polynomial Support Vector Machines (SVM) models were applied to the five test sites. The linear models consistently outperformed the polynomial SVM models and at Hornsey, Holderness and Dunwich, the best performing model was the linear SVM model with a regularisation parameter value of one (Figure 3.10 (*b*) – (*d*)). At Porthallow, the linear SVM model with regularisation parameter value of two outperformed all others (Figure 3.10 (*a*)). All outputs produced by the eight SVM models (four linear, four polynomial) at each site have been produced in Supplemental Materials A. The output of the best performing SVM model at each site for the four abovementioned sites is shown in Figure 3.10 (*a*) – (*d*).

The polylines produced at the same four sites by contouring the SVM outputs were compared to the manually digitised lines. SVM performed the best at Dunwich, (RMSE = 5.78 m, MAE = 4.23 m; Table 3.4). SVM performance was substantially worse at Hornsey compared with the other three sites (RMSE = 61.99 m, MAE = 31.47 m; Table 3.4). At Hornsey, the SVM prediction was less than 20 m from the manually digitised line at more than 80% of the transects. However, very large seawards error values were detected along some transects pertaining to exposed muddy substrate seawards of the coastal vegetation edge (Figure 3.10 (d); Figure 3.14 (*l*)). At all sites, the SVM models primarily predicted the coastal vegetation edge to be seawards of the manually digitised line, and at Holderness, the SVM prediction was seawards of the manually digitised line at every transect (Figure 3.14 (*c*), (*f*), (*i*) and (*l*)).

At Blakeney, no SVM model produced an output from which a vector polyline of the vegetation edge could be produced. With the exception of the linear model with regularisation parameter of 1, all models predicted a very small proportion (less than 5 %) of the entire image to be landwards of the coastal vegetation edge. The small number of pixels which were predicted to be landwards of the vegetation edge primarily corresponded with agricultural fields instead of the barrier island vegetation (see Supplementary Materials A).

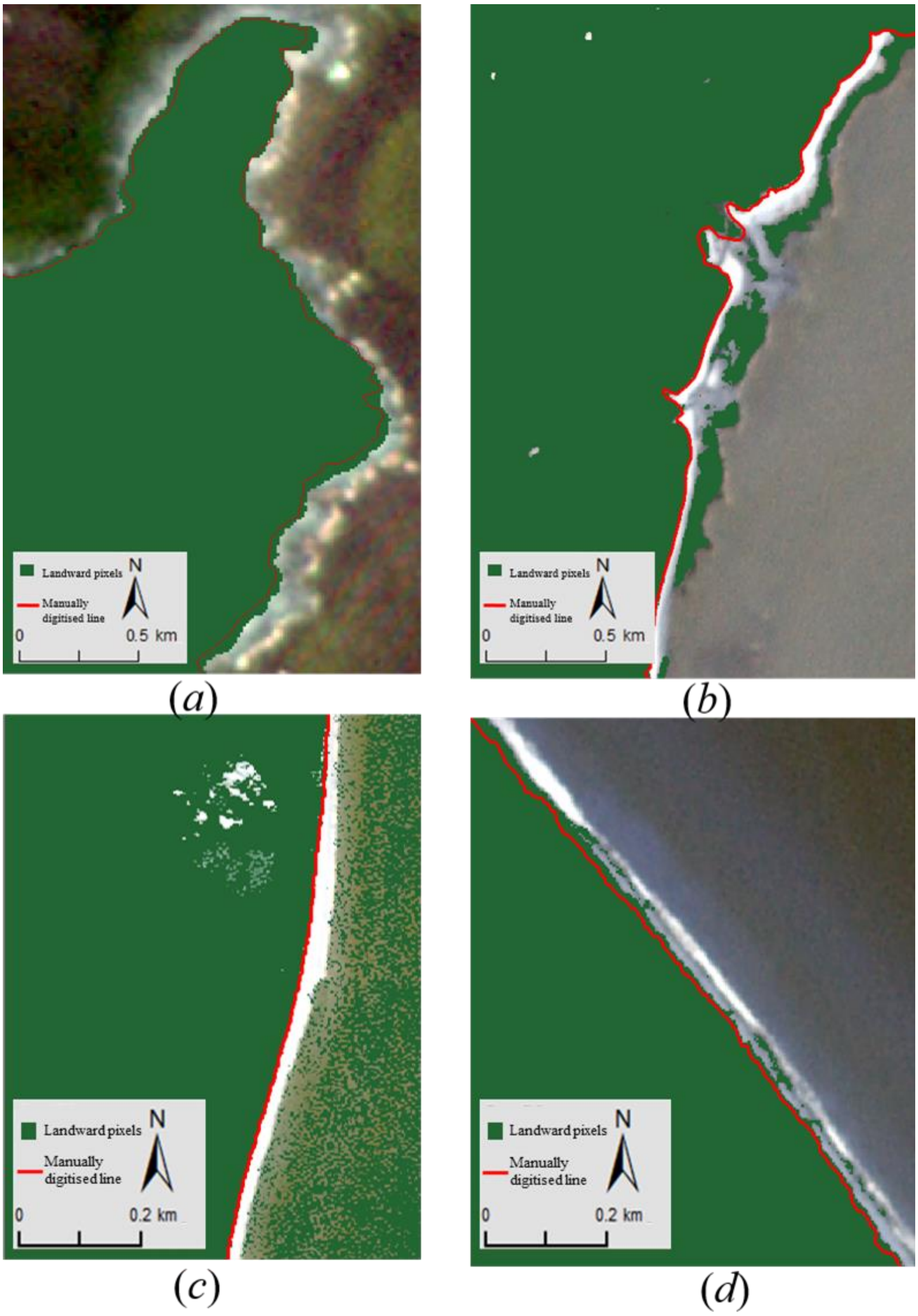


Figure 3.10: Outputs produced by the best performing SVM model at (*a*) Porthallow, (*b*) Hornsey, (*c*) Dunwich and (*d*) Holderness. Green pixels correspond to locations predicted to be landwards of the coastal vegetation edge. The red line represents the manually digitised line.

3.3.4. CoasTool

3.3.4.1. Row-wise NDVI pixel pattern

The cross-shore pattern in NDVI pixel values at Dunwich was determined by plotting the NDVI values for each row separately (Figure 3.11). For every cross-shore transect, NDVI values were negative in water, increased steeply at the waterline, plateaued near 0.0 across the beach before rising sharply again at the vegetation line. Cross-shore transects demonstrated that the vegetation line was situated at the transition between a set of adjacent pixels with relatively consistent, near-zero, NDVI values and a set of pixels whose NDVI values increased rapidly when traversing inland.

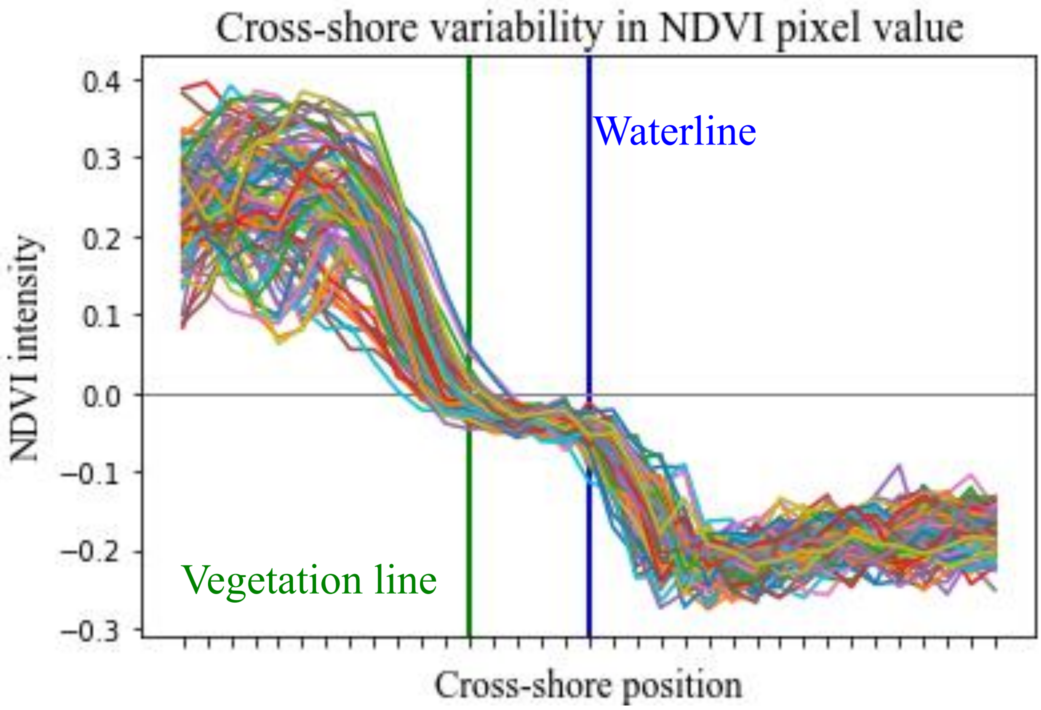


Figure 3.11: Cross-shore variability in NDVI pixel values for 100 rows of the NDVI image of Dunwich, Suffolk, UK. Land pixels are to the left of the image and sea pixels to the right. Each line represents a separate row in the image. The vertical blue and green lines represent the approximate water and vegetation line locations respectively. Each tick on the x-axis corresponds to a pixel.

Cross-shore variability in pixel values after passing the two kernels over the NDVI image and applying Equation (7) is shown in Figure 3.12. The application of Equation (7) leads to the generation of a small peak in pixel values close to the waterline, and a more distinct peak in pixel values at the vegetation line. The values of pixels inland of the vegetation line are approximately equivalent to the value of pixels seawards of the water line (Figure 3.12).

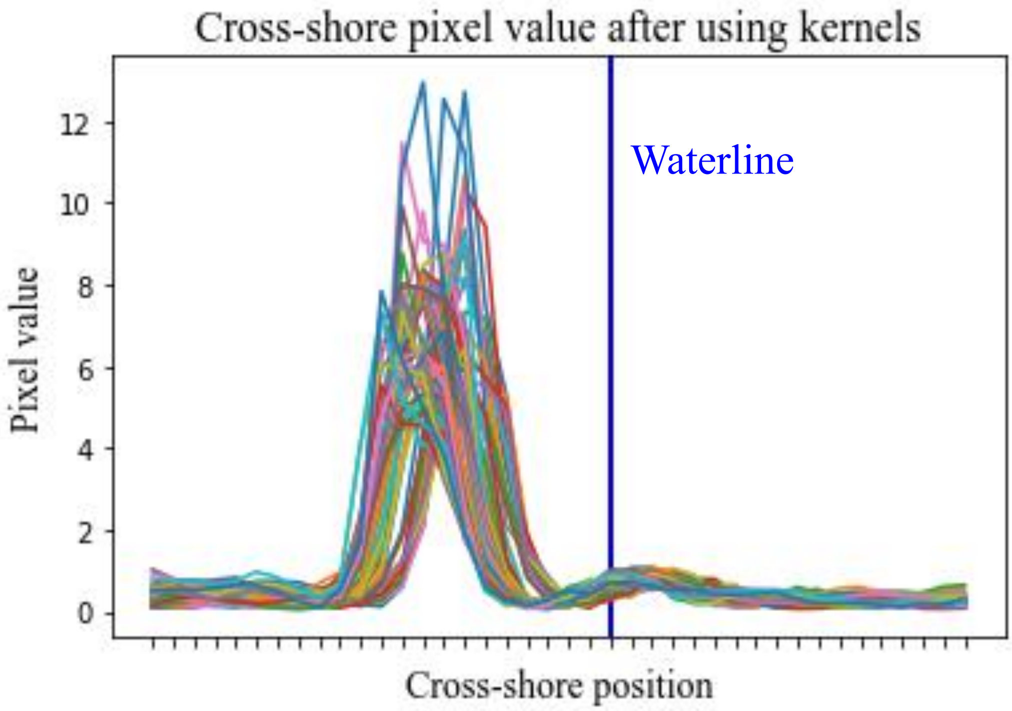


Figure 3.12: Cross-shore variability in pixel values at Dunwich, Suffolk, UK after passing the two 1-D kernels over the NDVI image and applying Equation (7). The vertical blue line denotes the approximate location of the water line and all pixels to the right of this line are seaward of the waterline. The major peak correspond to the approximate location of the vegetation line. All pixels to the left of the peak are landwards of the vegetation line peaks. The peaks are not aligned due to differences in beach width. Each tick on the x-axis corresponds to a pixel.

3.3.4.2. CoasTool performance

At all five sites, continuous vegetation lines were produced by CoasTool (Figure 3.13). RMSE and MAE values generated by CoasTool are shown in Table 4. Errors were lowest at Dunwich and Holderness (Dunwich: RMSE = 8.09 m, MAE = 6.61 m; Holderness:

RMSE = 4.75 m, MAE = 3.80 m). Errors were larger for the other three sites (Hornsey, Porthallow, Blakeney), being greatest at Blakeney (RMSE = 39.09 m, MAE = 23.26 m).

When comparing RMSE and MAE values across the three methods, CoasTool performed best at Hornsey and Blakeney, NDVI threshold contours performed the best at Holderness, and SVM outperformed the other two methods at Porthallow and Dunwich (Table 4). Errors produced by CoasTool were similar, but slightly larger, at Dunwich and Holderness than the errors produced by the 0.0 NDVI contour and SVM (CoasTool RMSE at Holderness = 4.75 m, 0.0 NDVI threshold contour RMSE = 2.31 m, Table 4). However, CoasTool errors at Blakeney and Hornsey were substantially lower than those produced using NDVI threshold contours and SVM (e.g. Hornsey CoasTool MAE = 17.33 m, 0.0 NDVI threshold MAE = 25.34 m, SVM MAE= 31.48 m; Supplementary Materials A).

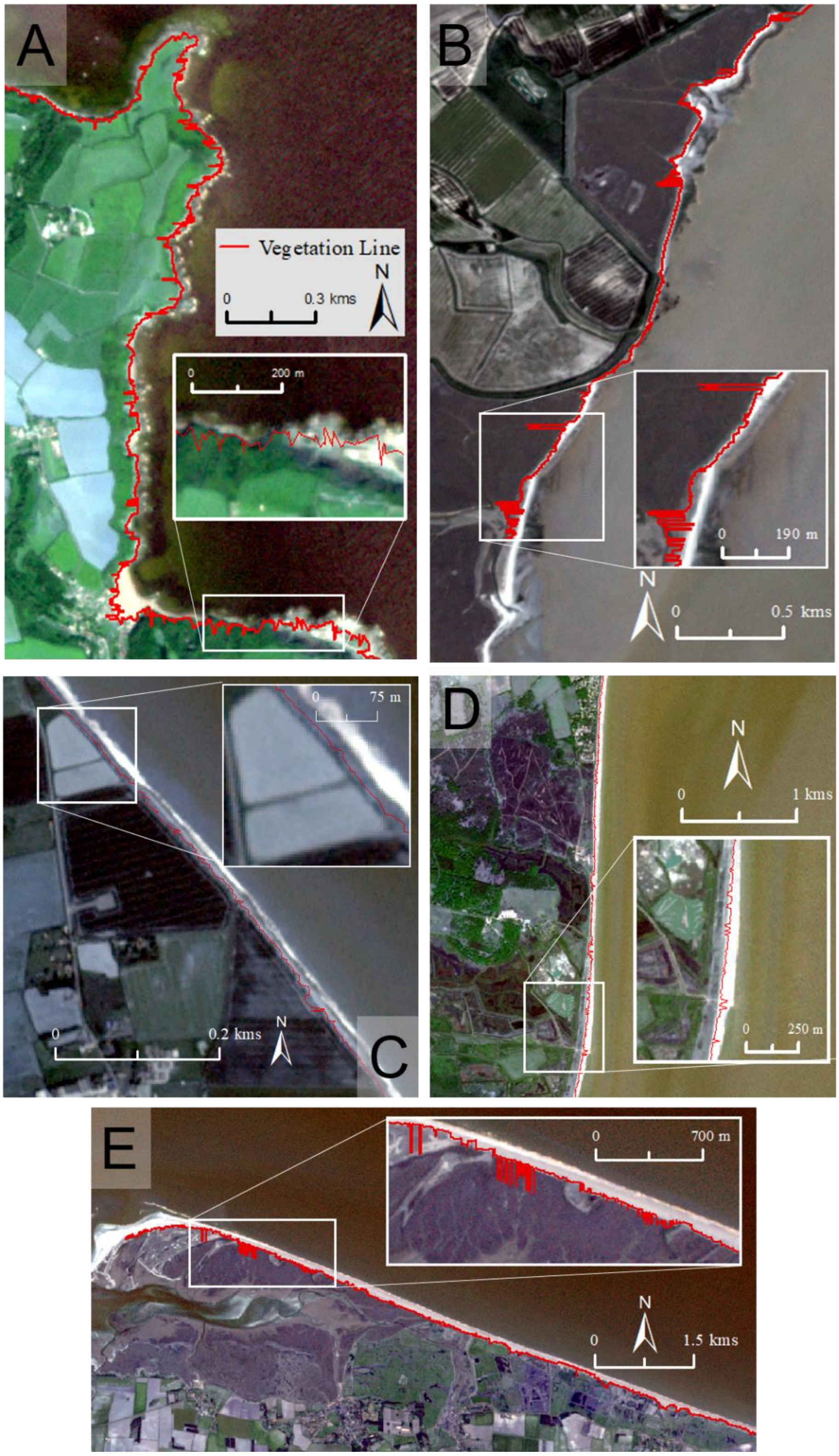


Figure 3.13: Vectorised CoasTool outputs at the five test sites (*a*) Porthallow, (*b*) Hornsey, (*c*) Holderness, (*d*) Dunwich and (*e*) Blakeney.

To further identify the scale and direction of the errors produced by CoasTool, NDVI threshold contours and SVM, histograms of error values for each site using each method are shown in Figure 3.14. The largest standard deviation (σ) value produced by the CoasTool is 37.95 m compared with 155.40 m by NDVI thresholds and 49.49 m by SVM. Very large σ values (greater than 90 m) were recorded by the NDVI threshold contours at Hornsey, Blakeney and Porthallow, corresponding to large negative (landward) errors in NDVI threshold contour location. SVM contour σ values were less than 11 m at three out of five sites, with most errors being seawards (positively) skewed (Figure 3.14). All three tools performed the best at Dunwich and Holderness, with MAE less than 9.97 m and σ less than 7.01 m values produced by each tool (Table 3.4, Figure 3.14).

Table 3.4: Comparison of CoasTool, NDVI threshold contours and SVM contour RMSE and MAE values at the five test sites. The best performing NDVI threshold contour is used at each site. The results of the best method at each site are highlighted in green.

	CoasTool		NDVI contour		SVM	
	RMSE (m)	MAE (m)	RMSE (m)	MAE (m)	RMSE (m)	MAE (m)
Dunwich	8.10	6.61	5.78	4.23	4.06	3.33
Holderness	4.75	3.80	2.31	2.13	11.62	9.97
Porthallow	23.27	14.83	108.43	46.22	11.15	7.32
Hornsey	34.63	17.33	92.76	25.34	62.00	31.48
Cromer	39.10	23.27	186.21	107.61	-	-

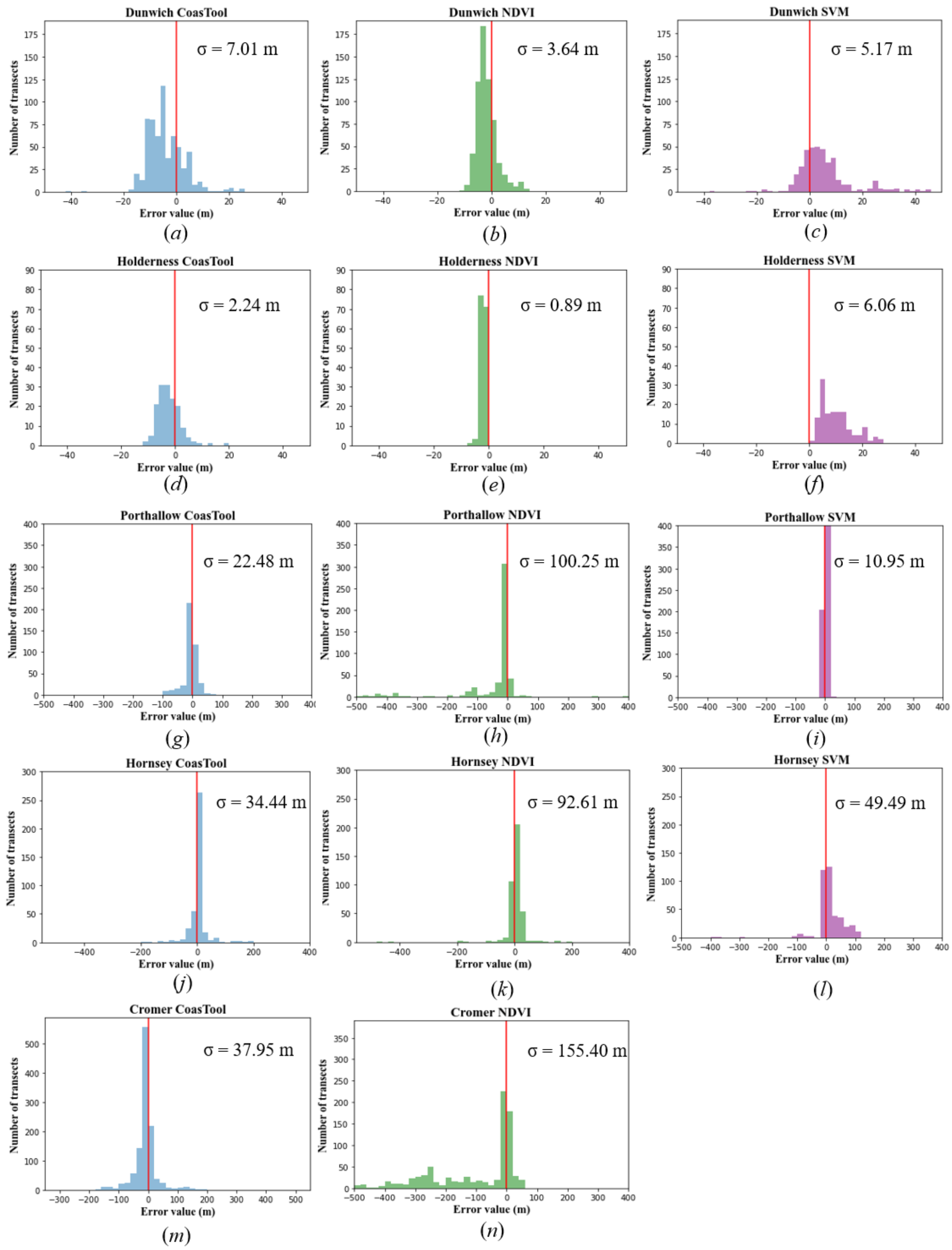


Figure 3.14: Histograms of error values produced by CoasTool, NDVI threshold contours and SVM. Positive and negative values correspond to transects where the tool outputs are seawards or landwards respectively of the manually digitised line. The red vertical lines denote a value of 0 m, where the manually digitised line and tool outputs overlap. Standard deviation (σ) values are provided for reference. Each site has a different scale on the x and y-axis, but every tool uses the same value range for both axes at each site. Note: No vectorised coastal vegetation edge could be produced at Blakeney.

3.4. Discussion

By considering the relative performance of different edge detection tools, more informed decisions can be made about the best tool to use in future studies. This chapter has described and applied a new edge detection tool, CoasTool, for the automatic extraction of the coastal vegetation line. The performance of CoasTool was compared to NDVI threshold contours, SVM classification and well-established edge detection operators. All tools accurately detected the coastal vegetation edge in multiple test images, generating RMSE less than 11 m or two image pixels. However, every tool produced RMSE greater than 35 m at one or more sites (Table 3.4).It is necessary to discuss the reasons for this varied performance before suggesting an alternative tool, convolutional neural networks.

The performance of all tools was assessed by comparing their output to manually digitised vegetation lines. Direct calculation of errors derived from manual digitisation was not possible because RTK-GPS measurements could not be taken from all five study sites. In a comparable study, Li et al., (2001) calculated mean positional errors of ± 8.5 m when manually digitising 4 m resolution IKONOS imagery. In Chapter 4, manually digitised vegetation lines are compared to ground-referenced measurements at three sites in Suffolkto derive a maximum RMSE of 4.13 m. The errors derived from manually digitising vegetation lines in this Chapter is therefore estimated to be 4.13 – 8.5 m.

3.4.1. CoasTool performance

This chapter applied CoasTool to five sites, each representing a distinct coastal setting. CoasTool uniquely considers the proximity of the coastal vegetation line to the land-sea interface and at all five sites CoasTool produced a continuous vegetation line with a Mean Absolute Error (MAE) less than 23.3 m. CoasTool was also able to take into account local variations in NDVI pixel values, instead of simply extracting pixels above or below a particular value, as in NDVI threshold contouring. CoasTool outperformed NDVI threshold and SVM contours at the two sites with the highest shoreline curvature: Blakeney, and Hornsey (Table 3.4) and also outperformed NDVI threshold contours at Porthallow. At Holderness and Dunwich, CoasTool performance was slightly worse but comparable to the NDVI threshold and SVM techniques. Although earlier studies have quantified errors derived from manually digitising coastal vegetation lines (Hapke et al., 2010; Leatherman, 2003; Li et al., 2001; Theiler et al., 2013), no study could be found which has quantified coastal vegetation edge positional errors using an automated technique. Error values derived from extracting the coastal waterline from remote sensing imagery with similar resolutions have been shown to range between 7.3 – 12.7 m using NDWI thresholding (Kuleli et al., 2011; Hagenaars et al., 2018; Vos et al., 2019a) and between 4.7 – 7.4 m using unsupervised classification (Garcia-Rubio et al., 2015). The results in this chapter are therefore comparable to those derived in other studies.

Along the relatively straight stretches of shoreline at Dunwich and Holderness, CoasTool produced small errors, less than two image pixels in size (Holderness MAE = 3.80 m; Dunwich MAE = 8.09 m). At these sites, the relatively consistent angle of the shoreline across the entire image enabled the 1-D kernels to detect the cross-shore patterns in NDVI values that this chapter has discovered (Figure 3.11); namely that the coastal vegetation line is situated at the interface between adjacent pixels with near-zero NDVI values and pixels which rapidly increase in NDVI value further inland. In contrast, along more complex coastlines exhibiting a higher curvature, the pattern of pixel values described in Figure 3.11 was not consistently found. This led to CoasTool performing worse at the three sites containing a more complex shoreline, most notably Porthallow (MAE = 14.83 m, RMSE = 23.37 m, Figure 3.13 (*a*)).

CoasTool was also susceptible to elevating the value of pixels not corresponding to the vegetation line but which were situated in a row of pixels exhibiting the same cross-shore pattern identified in Figure 3.11. This led to pixels corresponding to the drift line at Dunwich and boundaries between sand and rock at Porthallow being erroneously detected as the vegetation line (Figure 3.13 (*a*) and 3.13 (*d*)). Erratic fluctuations in vegetation line position were detected at some locations, e.g. at Hornsey and Blakeney (see insets Figure 3.13 (*b*) and 3.13 (*e*)), resulting in CoasTool output pixels being greater than 100 m inland of the manually digitised vegetation line along some transects (Figure 3.14 (*j*) and 3.14 (*m*)). This is attributable to CoasTool detecting a more rapid increase in NDVI pixel values inland of the most seaward coastal vegetation line. This aligns with studies looking at cross-shore gradients in vegetated land cover in salt marsh habitats, where the most abrupt increase in vegetation cover (and correspondingly pixel NDVI values) is situated inland of the most seaward salt-tolerant species (Feilhauer et al., 2020; Unberath et al., 2019).

3.4.2. SVM performance

SVM outperformed CoasTool and NDVI threshold contouring at two of the five sites (Dunwich and Porthallow), although a very large number of pixels seawards of the vegetation line were erroneously predicted to be landwards by the SVM (Figure 3.10 (*c*)). This high performance shows promise in using supervised machine learning methods to detect the coastal vegetation edge, attributed to their ability to iteratively learn the relationship between spectral values and class, use all spectral bands and be trained using a relatively small dataset (Elnabwy et al., 2020; Mountrakis et al., 2011). By comparison, the other three methods were restricted to using a single-band image as input.

SVM performed the worst at Hornsey where exposed muddy substrate was erroneously predicted to be landwards of the vegetation edge (Figure 3.10 (*b*)) and at Blakeney where no vectorised vegetation line could be produced. The poor performance of SVM at these sites is attributed the very graded coastal vegetation boundary at Hornsey and Blakeney, emphasised by the large distance between NDVI threshold contours at these sites. This corresponds to the results of Choung and Jo (2017) where the performance of SVM at predicting the coastal waterline position were lowest where the waterline boundary was less distinct, primarily corresponding to shaded locations and locations with shallow beach gradient. Hornsey and Blakeney also represent coastal settings very distinct from those contained within the training dataset. Blakeney is a gravel spit with adjoining salt marshes (Hardy, 1964, Pollard et al., 2020) and Hornsey contains salt marsh vegetation species with fronting muddy substrates (Scrimshaw et al., 1996). The spectral differences in the vegetation and substrate may explain the poorer performance of the SVM models at these two sites. This highlights the fact that alongside considering the size of the training dataset, the training images need to encompass a range of coastal settings if SVM models, or indeed other supervised machine learning tools, are to generalise to accurately predict the coastal vegetation edge location in multiple locations.

3.4.3. NDVI threshold contour performance

The position of the vegetation line produced by NDVI contours was sensitive to the threshold value used. At Hornsey and Blakeney Point, the salt marsh and barrier island vegetation commonly had NDVI values below 0.05. At Hornsey, the 0.1, 0.075 and 0.05 NDVI threshold contours detected the back barrier of the salt marsh. At Blakeney, large stretches of the westerly edge of the barrier island could only be detected using the 0.0 NDVI threshold contour (Figure 3.5 (*b*) and 3.5 (*e*)). Whilst the 0.0 NDVI threshold contour produced the smallest errors at four out of the five sites, the 0.0 contour produced the largest errors at Porthallow (Table 3.2). At this site, the 0.0 contour was situated seawards of the waterline along some transects, likely due to the presence of macroalgae or other substrates in the water column elevating the corresponding NDVI pixel values (Figure 3.5 (*a*)). This highlights how NDVI threshold contour position is sensitive to biotic factors including plant species, composition and phenology and environmental factors including soil moisture content and substrate composition (Belluco et al., 2006; Gandhi et al., 2015; Rahman et al., 2011).

NDVI threshold contours are also unable to overcome small along-shore breaks in the vegetation line, often leading to substantial, unnatural migrations of the NDVI contour inland (Figure 3.14 (*h*) and (*k*)). At Holderness, where a ploughed agricultural field was detected, the 0.05, 0.075 and 0.1 contours followed the vegetated margins of the field, causing the NDVI contours to be situated up to 108.3 m landwards of the manually digitised line (Figure 3.5 (*c*)). These largescale inland migrations of higher threshold NDVI contours explain the larger differences between RMSE and MAE values for 0.1 contours compared with 0.0 (Table 3.2). Higher weights are placed on large error values when calculating RMSE compared with MAE (Chai and Draxler, 2014). Larger differences in the RMSE and MAE values produced by higher threshold contours demonstrates how error values are low for the majority of transects but that a low number of transects contain large errors. The large variability in NDVI threshold contour position due to gaps in the vegetation line, and sensitivity in the threshold value used, precludes the fully automated use of NDVI threshold contours to detect the coastal vegetation line position in remote sensing imagery.

3.4.4. Edge detection operator performance

This chapter has also demonstrated that edge detection operators are not suitable for exclusively detecting the coastal vegetation edge in remote sensing imagery. When a non-buffered vegetation line was used, the highest recorded P_A value recorded was 0.13, meaning only 13% of pixels corresponding to the manually digitised vegetation line were identified by the kernel operators as being an edge location (Table 3.3). The edge detection operators did not perform consistently better when applied to either the NDVI or greyscale image. Edge detection operator performance was better when applied to the NDVI image at Holderness (Figure 3.6) and Porthallow (Figure 3.7) but performance was better when applied to the greyscale images at the other three sites (Table 3.3). It was anticipated that edge detection operator performance would be better on the NDVI images because the key function of the NDVI function is to distinguish vegetated from non-vegetation land covers (McFeeters et al., 1996). It is suggested that the graded nature of the coastal vegetated boundary reduced the performance of applying the edge detection operators on some images (Feilhauer et al., 2020; Unberath et al., 2019). The most consistent performance by the kernels applied to both single-band images was at Porthallow (Sobel NDVI: $U_A = 0.86$, $P_A = 0.31$, Roberts greyscale: $U_A = 0.49$, $P_A = 0.15$). This is attributed to the relatively steep gradient in both NDVI and greyscale pixel intensity values at this site due to the close proximity of the vegetation and water line (Figure 3.7 (*a*) – (*d*)). These findings align with other studies which have attributed the poor performance of the edge detection operators to the lack of consistent, sufficient contrast in pixel intensity values at the waterline (Liu and Jezek, 2004).

Whilst kernel performance was high at some sites, the inconsistency in performance depending on the single-band input image used precludes the fully automated use of the edge detection kernels to exclusively detect the coastal vegetation edge. This Chapter's results support previous assertions that the edge detector operators locate the positions of the greatest rate of change in pixel greyscale intensity (Liu and Jezek, 2004). The objective of this study, however, was not to find the most abrupt edge but the semantically correct edge i.e. exclusively the coastal vegetation edge. The key limitation of all the edge detection operators at every site is that they also detected other, irrelevant boundaries within the image, including the coastal waterline and inland boundaries such as field margins and anthropogenic land covers. Sobel and Roberts' approaches consistently outperformed Laplacian methods. This was attributed to the line of zeros contained within the Sobel and Roberts kernels which makes them more suitable for finding linear boundary features (Figure 3.6 (*b*) & (*d*), 3.7 (*b*) & (*d*), 3.8 (*b*) & (*d*) and 3.9 (*b*) & (*d*); Al-Amri et al., 2010). Conversely, the Laplacian kernel is better suited to finding individual pixels with substantially different values to pixels in all other directions, leading to the 'speckled' appearance of many of the Laplacian kernel outputs (Figure 3.6 (*c*), 3.7 (*c*), 3.8 (*c*), and 3.9 (*c*)).

No automated way to derive an image-specific threshold value was found to filter the outputs of the edge-detection operator outputs. Instead, this study arbitrarily retained only the 5% of pixels with the highest values. Retaining a smaller percentage of pixels would have reduced the number of false positives, potentially increasing U_A values, but this would have also increased the number of false negatives, potentially reducing P_A values. The lower 95% of pixels were removed because to make the analysis in this Chapter analogous to the post-processing steps applied to the outputs of convolution neural network (see Chapter 4), where the position of pixels predicted to be the coastal vegetation edge with a value ≥ 0.95 were compared.

3.4.5. Further research requirements

All four tools used in this chapter represented the vegetation edge as a discrete line or single pixel. This Chapter questions whether this is the most appropriate way to represent the coastal vegetation line. When applying Equation (3.7), CoasTool elevates the values of a cross-shore zone of pixels, rather than one individual pixel per row (Figure 12). Likewise, when using NDVI contours at Dunwich, where small errors were produced by all contours, there was still a maximum discrepancy of 37.6 m in NDVI contour position (Figure 3.5 (*d*)). The 0.0 contour was likely to have detected isolated, seaward clumps of pioneer vegetation species, whereas the 0.1 contour detected more landwards, continuous land cover (Figure 5 (*d*)). This highlights how the coastal vegetation edge is not an abrupt boundary but rather a zone where vegetation cover density gradually increases when traversing cross-shore inland from the exposed beach or other intertidal zone.

The 'fuzziness' of coastal vegetation boundaries has led to the application of soft-classification methods which provide a probabilistic rather than absolute value that a pixel corresponds to different coastal vegetation species (Berhane et al., 2018; Wen et al., 2020). Probabilistic classification better accounts for the continuum in vegetated land cover, typically characterised by an increasing density in vegetation cover when traversing inland

(Feilhauer et al., 2020). These soft-classification methods, however, are still unable to distinguish between coastal and more terrestrial vegetation species (Rahman et al., 2011).

Machine learning tools are a potential alternative for dealing with the fuzzy nature of the coastal vegetation boundary. Supervised machine learning tools can be provided with initial coastal remote sensing input images and corresponding binary edge/ non-edge maps (Kokkinos, 2015). This potentially provides the supervised machine learning tools with semantic information, required to detect the 'correct edge', instead of simply the locations with the greatest gradient in pixel greyscale intensity value (Chong and Jo, 2017; Kokkonos, 2015; Wan et al., 2019). A form of machine learning, Convolutional Neural Networks (CNN), has been used to derive probabilistic outputs that a particular pixel represents the position of field boundaries (Waldner and Diakogiannis, 2020); inland waterbody boundaries (Chen et al., 2018) and mangrove extent (Wan et al., 2019), but has not been applied to detect the coastal vegetation edge in remote sensing imagery.

Heterogeneity in the spectral properties of coastal vegetation also limited the performance of all tools used in this Chapter (Table 3.4). Attributes such as the satellite platform used, time of year, time of day and phenology all heavily influence the spectral properties of vegetation (Belluco et al., 2006; Rahman et al., 2011). When trained on a substantial number of images, supervised machine learning tools have been able to detect field boundaries in images captured by different satellite platforms at different azimuths (Waldner and Diakogiannis, 2020), and to detect the same inland vegetation margin in images captured during all seasons (Watkin and van Niekerk, 2019). The ability to consider that vegetation can take a range of spectral values is essential to deriving an automated tool which can consistently detect the coastal vegetation edge from multispectral remote sensing imagery (Feilhauer et al., 2020; Wan et al., 2019).

Irrespective of the shortcomings identified in all of the methods used in this Chapter, all methods were able to detect a continuous coastal vegetation edge, distinct from the land-water interface (Figure 3.5, 3.10 and 3.13). This is primarily due to the use of 3 – 5 m spatial resolution Planet imagery (Table 3.1). The ability to automatically and simultaneously extract both proxies from multispectral remote sensing imagery provides promise in further understanding coastal dynamics at annual to decadal timescales. The instantaneous waterline can identify very rapid shoreline responses to meteorological and hydrological forcing factors, whereas the vegetation line provides a longer-term representation of shoreline dynamics (Zarillo et al., 2008). Further investigation is required to understand whether further insights into coastal zone dynamics can be generated from simultaneously extracting these two shoreline proxies.

3.5. Conclusion

This Chapter has described and applied CoasTool to automatically detect the coastal vegetation line in remote sensing imagery. CoasTool has been shown to outperform NDVI threshold contours and edge detection operators along more complex coastlines with higher curvature. However, the three abovementioned tools detected irrelevant boundaries and detected the vegetation line as a discrete line or single pixel. This does not represent the graded nature of the coastal vegetation edge, although CoasTool and the edge detection kernels do tend to elevate a region of pixels near the vegetation edge, rather than an individual pixel (Figure 3.12). Further investigation is required into tools which may better represent the graded nature of the coastal vegetation edge.

A key challenge in detecting the coastal vegetation edge is that the most seaward vegetation can exhibit a range of pixel intensity values. CoasTool was able to identify locations where the NDVI values increased compared with the exposed intertidal substrates and did not need to find absolute NDVI values as in the thresholding methods. However, similar to the edge-detection operators, CoasTool lacked semantic information to sometimes differentiate between the seaward coastal vegetation edge and other locations with abrupt changes in NDVI value, and was susceptible to noise, for example if a drift line was present in the intertidal zone. In comparison, SVM outputs tended to be able to discard irrelevant inland boundaries (Figure 3.10), primarily attributed to the supervised nature of the tool, but was unable to generalise to a range of coastal settings. Further investigation into the ability of supervised machine learning methods to exclusively detect the coastal vegetation edge are required to identify whether they can address some of the limitations of the methods investigated in this Chapter.

Chapter 4. VEdge_Detector: Automated coastal vegetation edge detection using a convolutional neural network

4.1. Introduction

Convolutional Neural Networks (CNN) have recently received increased attention as a way to effectively detect edges in remote sensing imagery. This is in part because they simultaneously consider the value of the pixel of interest and neighbouring pixels (Kokkinos, 2016; Zhang et al. 2016). CNNs convolve kernels of different sizes over the raw input image. Smaller kernels (e.g. 3 × 3) capture detailed edge structures but suffer from high incidence of false positives (noise). Conversely, larger kernels detect only the most salient edges, generating blurred boundaries and missing localised detail. Optimal fusing of the outputs from different sized kernels subsequently identifies the most likely location of true edges and minimises noise by considering that edges will be in the same location irrespective of kernel size (Ren, 2008).

Holistically-Nested Edge Detection (HED) is an example of a CNN which progressively reduces image resolution, instead of increasing kernel size, to achieve multi-scale image convolution (Xie and Tu, 2015). The HED model architecture contains five separate sets of convolutional layers, all using 3 × 3 kernels, which are each separated by 2 × 2 maximum pooling layers to reduce image resolution. A side output layer is produced after every set of convolutional layers. The first side output contains local boundary detail but is susceptible to noise and false inland boundaries. Conversely, side output 5 only detects salient boundaries and is robust to image noise but the predicted coastal vegetation edge is blurred. These five side output layers are optimally fused to derive the final output, predicting the likelihood of each pixel being an edge (Xie and Tu, 2015; see Figure 4.1 for a graphical overview of HED architecture).

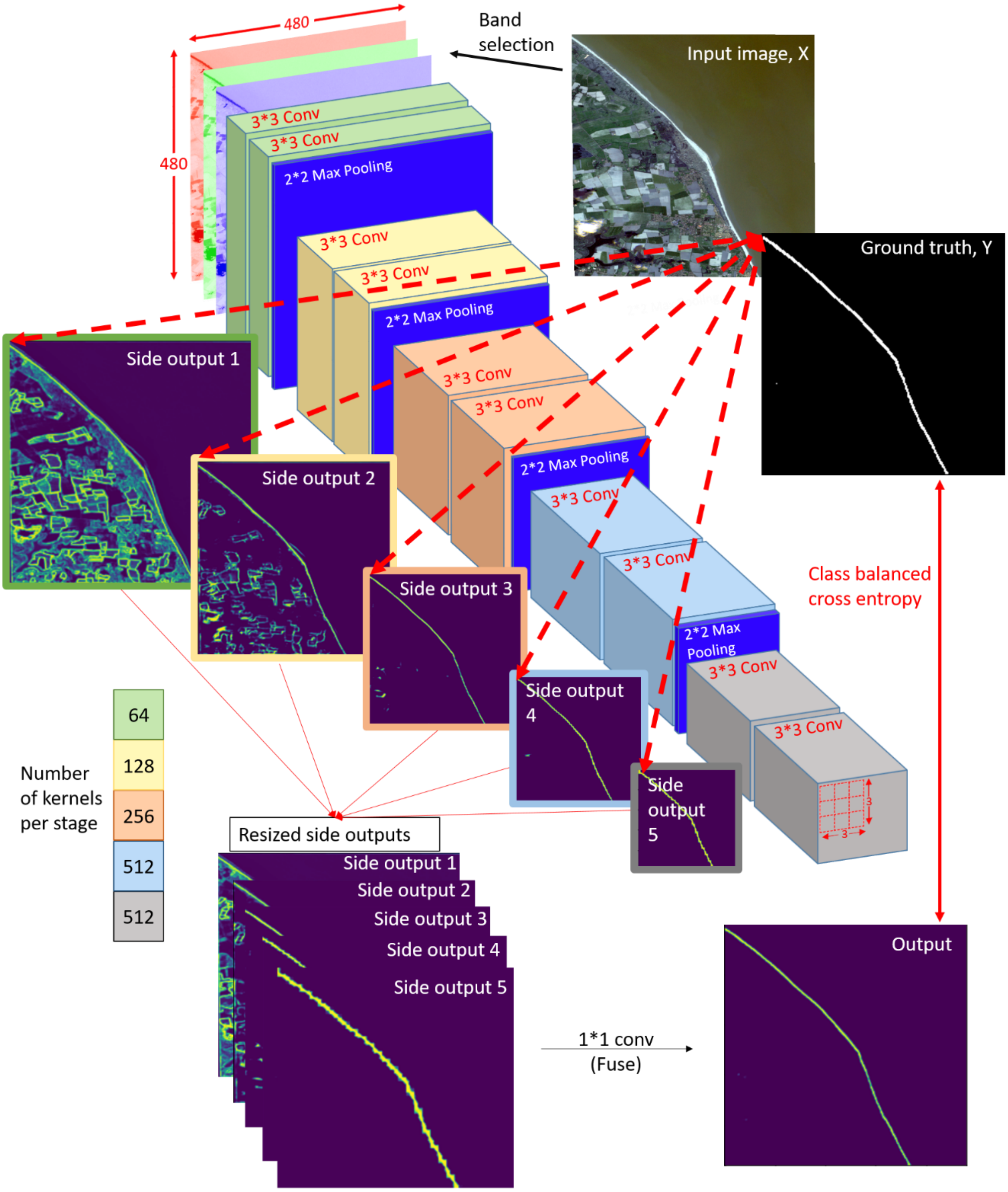


Figure 4.1: Holistically-Nested Edge Detection (HED) architecture. Three spectral bands from every satellite image are selected as HED input. Input images are fed through five distinct stages of image convolution, and between each stage a max pooling layer decreases image size. The squares to the bottom left of the image detail the number of convolution kernels at each stage. The side outputs are resized and optimally fused to generate the output.

Applications of CNN methods, including HED, to detect edges have recently increased in number, due to enhanced computer processing power and greater image availability to train the CNNs e.g., natural image datasets including the Berkeley segmentation dataset (Arbelaez et al. 2007) and ImageNet (Stanford Vision Lab, 2016). The Visual Geometry Group Network (VGGNet-16) model is a CNN with a very similar architecture to HED but contains no side outputs. The model was trained using the ImageNet dataset to detect all objects in natural Red Green Blue (RGB) images e.g., images of animals, humans and everyday items (Simonyan and Zisserman, 2015). Applications of HED to detect every object in natural images are widespread but remote sensing applications, where images contain more noise and a higher density of boundaries, remain highly limited. A key research gap is the retraining and fine tuning of these generalist edge detection CNNs to be able to differentiate between separate types of edge in remote sensing imagery and exclusively extract edges of interest.

To exclusively detect particular types of edge in remote sensing imagery, some studies have updated or fine-tuned the weights within pre-existing CNNs by retraining them with remote sensing image pairs. Richer Convolutional Features (RCF), which are CNNs with a similar architecture to HED, have been fine tuned to exclusively detect building boundaries in remote sensing imagery, achieving a higher accuracy than other generalist edge detection algorithms (Lu et al., 2018). In Lu et al.'s study, fine tuning was conducted by training the RCF on 1856 image pairs containing an urban scene and a binary image showing building edge and non-edge locations. Similarly, a U-Net neural network was retrained with Landsat imagery to predict glacial calving front locations (Mohajerani et al., 2019). Remote sensing applications of HED, or modified versions, have been used to detect field boundaries (X.Y. Liu et al., 2019) and to derive land cover classification (X.Y. Liu et al., 2019; Marmanis et al., 2018). H. Liu et al. (2019) modified the standard convolution structure of HED to detect shorelines in heavily urbanised Jiaozhou Bay, China. HED was reported to outperform Sobel and Canny Edge Detection (producer accuracy: Sobel = 0.66, Canny = 0.82, modified HED = 0.95) but no information was provided on the shoreline proxy used. Furthermore, this study trained HED using exclusively RGB spectral bands; further analysis is necessary to identify the optimum spectral band combination during HED training. These abovementioned studies highlight the potential of retraining a CNN to fine-tune its internal weights to exclusively detect a particular type of edge in remote sensing imagery. To date, this approach has not been applied to exclusively detect coastal vegetation edges from remote sensing imagery.

This chapter aims to train and apply a Holistically-Nested Edge Detection (HED) model to extract coastal vegetation lines. The objectives of the chapter are to: i) train a HED model using coastal remote sensing imagery, namely Planet 3 m and 5 m resolution imagery (PlanetScope and RapidEye); ii) assess the performance of HED in extracting the coastal vegetation line when trained using different combinations of spectral bands as input across a range of coastal settings (Winterton, Norfolk, UK; Perranuthnoe, Cornwall, UK; Bribie Island, Australia and Wijk-aan-Zee, The Netherlands); iii) compare the performance of a new edge detection tool, Vedge_Detector , against other experimental methods previously used to detect the coastal vegetation edge, namely ground-referenced measurements and manual digitisation of remote sensing and aerial imagery; and iv) incorporate the best performing HED model within VEdge_Detector to detect shoreline change from sequential images of Covehithe, Suffolk, UK between 2010 to 2020.

4.2. Materials and Methods

4.2.1. Remote sensing imagery data sources

A total of 78 Planet images (PlanetScope and RapidEye, with 3 and 5 m spatial resolution respectively) were selected for HED training (Planet Team, 2017). Ortho Scene product level imagery was chosen, meaning Planet had orthorectified and radiometrically corrected images prior to image download. Locations were chosen to encompass a diverse range of geomorphic landforms, tidal ranges and vegetation types (see Supplementary Material B). Training image sizes ranged from 6.3 km^2 to 1557.5 km^2 and images were selected from all years when Planet imagery was available in the period 2010 to 2020. Multiple images were collected from each location to ensure the training dataset contained scenes captured at different tidal stages. This ensured that multiple images of the same shoreline, but with different beach widths, were contained in the training dataset.

4.2.2. Holistically-Nested Edge Detection (HED) training

All steps taken in this study were separated into three stages: HED training using coastal remote sensing imagery; validation of the trained HED models; and digital shoreline change analysis using the best performing HED model. The training and validation stages determined the optimal combination of remote sensing spectral bands to train the HED model whilst keeping the HED model architecture constant. The best performing HED model became the VEdge_Detector tool, developed to extract vegetation lines in the shoreline change stage. Figure 4.2 provides a graphical overview of the three analytical stages.

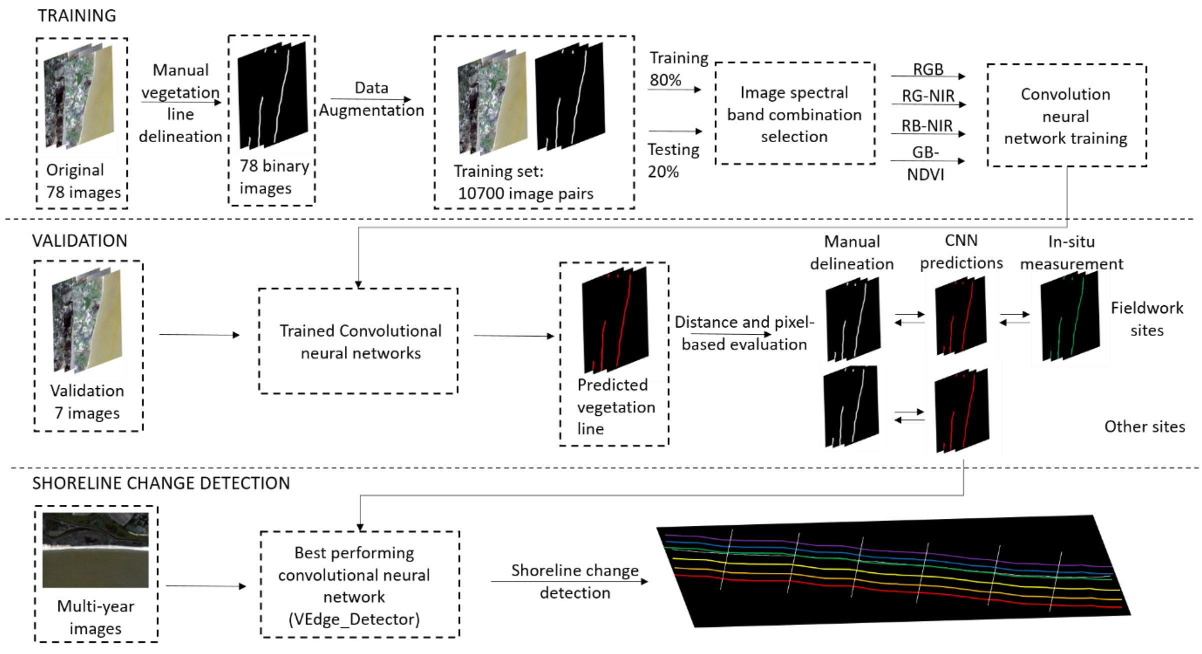


Figure 4.2: Overview of the three stages of VEdge_Detector training and application. Four Holistically Nested Edge Detection (HED) models were independently trained using different spectral band combinations (training). The performance of each HED model was evaluated using a separate image set (validation). The best performing HED model, trained on images with spectral band combination red, green, near-infrared (RG-NIR), formed the VEdge_Detector tool. This tool detected the vegetation line position from multiple images of the same shoreline captured over a 10-year period (shoreline change detection).

4.2.2.1. Manual digitisation of the vegetation line

To generate the training dataset, vegetation lines were manually digitised from all 78 training images in ArcGIS 10.5.1. The image NDVI was overlaid at 70% transparency to aid visual vegetation line identification. Where vegetation lines were interrupted, the seaward extent of inland waterbodies or urban areas was used. Vegetation line shapefiles were converted into binary raster edge maps (binary images), with edge pixel values set to 1 and non-edge pixels to 0. Image pairs were subsequently established, containing the original image and the binary image.

4.2.2.2 Data Augmentation

A large number of images are required during HED training to refine the internal weights within the HED model. Manual digitisation of this number of images would be too time consuming; therefore data augmentation was used to substantially increase training data size, from 78 to 10 700 image pairs. Larger images were cropped to size 480 × 480 pixels (the default image size used by the HED architecture) at multiple locations. The uncropped larger images also formed part of the training dataset but were resized to 480 × 480 pixels prior to HED training. Image pairs were flipped vertically, rotated by 90, 180 and 270 degrees and subject to the introduction of Gaussian noise (Figure 4.3). Gaussian noise was not added to the binary images. Images were rotated around five different points of origin. Image pairs not containing any vegetation line after rotation were automatically discarded. All image pairs were shuffled and randomly assigned into training (80%) and testing (20%) sets prior to CNN input. The proportion of land cover in each image varied from 2% to 98%.

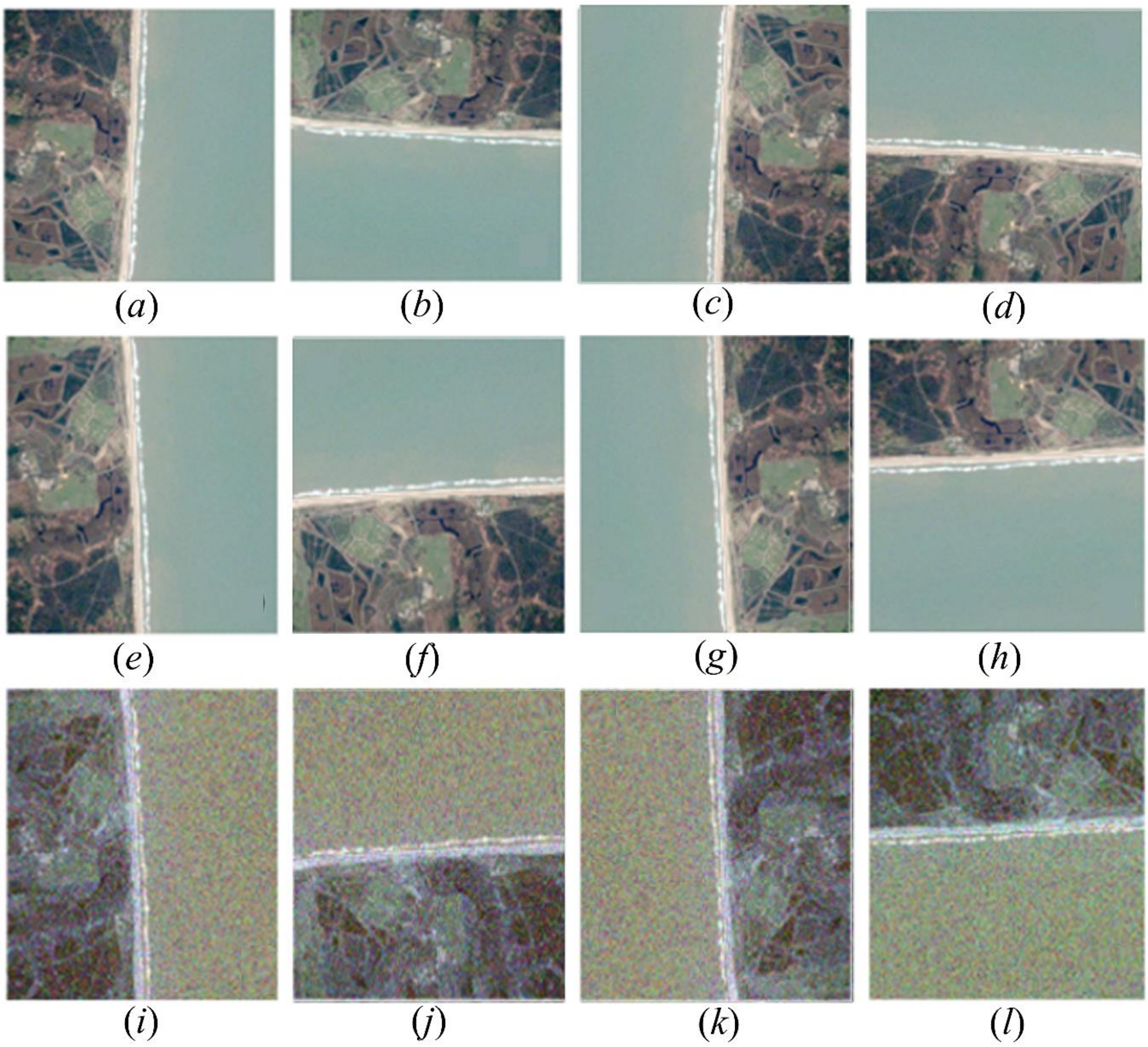


Figure 4.3: Transformations used in data augmentation (*a*) original image (cropped to 480 × 480 pixel size), (*b*) – (*d*) original image rotated by 90°, 180° and 270°, (*e*) original image flipped vertically, (*f*) – (*h*) flipped image rotated by 90°, 180° and 270°, (*i*) – (*l*) Gaussian noise added to the flipped images. Transformations (*b*) – (*h*) were simultaneously conducted on the binary images.

4.2.2.3. Holistically-Nested Edge Detection (HED) training

HED training was conducted to modify the model's internal weights to increase the model's ability to exclusively detect coastal vegetation edges. To speed up HED training, non-zero weights were initialised prior to training commencement. This study initialised the internal weights contained within the VGGNet-16 architecture prior to training. The weights contained within the VGGNet-16 architecture were derived from training the model on 1.2 million natural images to detect everyday objects e.g. animals, people and urban features. Using the weights contained within the VGGNet-16 architecture increased the speed of HED training compared to using randomly assigned weights. The key difference between the architecture in VGGNet-16 and HED is that HED contains side outputs. The side outputs enable deep supervision, whereby every side output is compared to the binary image to calculate loss. By comparison, in VGGNet-16 only the final output is compared to the binary image. Deep supervision guides the neural network to detect transparent objects, i.e., to only detect the edges of objects at a per-pixel level rather than the entirety of an object (Xie and Tu, 2015).

To substantiate the assertion that the default weights in the VGGNet-16 architecture were not suitable for detecting exclusively coastal vegetation edges, a HED model containing the default VGGNet-16 weights was used to predict the coastal vegetation edge in an image of Winterton, Norfolk, UK. This HED model failed to detect the coastal vegetation line and instead detected the water line and other inland boundaries (e.g. roads and field edges). This was attributed to the weights in the VGGNet-16 architecture originally being trained to classify all objects in a natural RGB image whereas the objective of this study was to exclusively extract the vegetation line in remote sensing imagery and discard other boundaries. This need reinforced the necessity to retrain the HED model to refine the model weights, using the image pairs derived through manual digitisation and data augmentation.

During every epoch of HED training, the internal weights in the HED model were used to predict the coastal vegetation line position from the raw image. The class-balanced cross entropy loss function was used to calculate the difference, or loss, between the predicted vegetation line position and the binary image. The loss function was class-balanced to account for the large imbalance between edge and non-edge pixels; the vast majority of pixels in every image were non-edge. To prevent the HED model from achieving very accurate results if it predicted all pixels to be non-edge, a scaling factor was used. This was calculated by determining the proportion of edge to non-edge pixels in each image. This scaling factor ensured that the HED model was penalised proportionately more for predicting a false negative (predicting an edge pixel to be a non-edge) than a false positive (predicting a non-edge pixel to be an edge).

HED model training was implemented in Python’s Keras library with Tensorflow backend.

The code for the training of HED was modified from Liu (2018) to enable input of 16 bit Planet imagery; selection of the desired image band combination; and the calculation of NDVI. The HED model was run in parallel on four Tesla P100-PCIE-16GB GPUs for 1000 epochs, with a running time of seven hours 45 minutes per spectral band combination. The VEdge_Detector tool, instructions and input image specifications are freely available from GitHub (github.com/MartinSJRogers).

4.2.3. Validation

The HED model performance was validated by predicting the vegetation line location in seven images not previously seen by the model. All output prediction pixel values ranged between 0 to 1, representing the range in HED confidence that the pixel represented the vegetation line. Confidence contours were used to determine where ground referenced measurements were located in relation to predicted vegetation line confidence curves. HED outputs were accordingly contoured at 0.1 intervals between 0.05 to 0.95 for subsequent model evaluation through comparison with ground-referenced measurements. All contours had a landward and seaward line (see Figure 4.4 for a demonstration of the vegetation line contours produced).

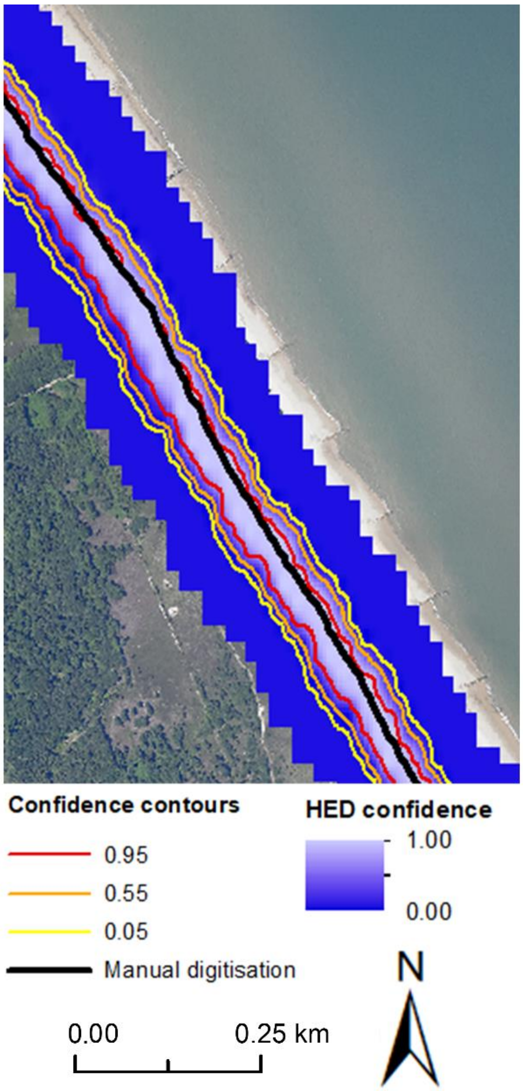


Figure 4.4: Example of 0.05 (yellow), 0.55 (orange) and 0.95 (red) confidence contours produced by VEdge_Detector at Winterton, UK. The confidence contours were generated from the raw VEdge_Detector output, which is overlaid as the blue colour ramp. Light and dark blue pixels represent the locations predicted as being an edge pixel with a high and low confidence respectively. The manually digitised vegetation line (black) is displayed for visual comparison. Land and sea are found to the left and right of the image respectively. Aerial imagery, provided by the Environment Agency with 40 cm resolution, is used as a backdrop (Environment Agency, 2020a).

Distance and pixel-based evaluation metrics were used to determine the best performing HED model. Distance-based evaluation of HED performance was conducted by comparing: i) HED model prediction contours (confidence contours) with ground-referenced measurements of vegetation line location; ii) confidence contours to a manually digitised vegetation line of the same image; and iii) ground-referenced measurements to manual digitisation.

The ArcMap plugin Digital Shoreline Analysis System (DSAS; (Thieler et al., 2009; USGS, 2018)) v5.0 was used in ArcGIS 10.5.1 to calculate distance between shorelines for comparators i), ii) and iii). Distance calculations were made on transects generated at 10 m alongshore intervals, orthogonal to the dominant shoreline orientation. To reduce transect crossing on sinuous coastlines, each transect was drawn orthogonal to a smoothed baseline, generated approximately 200 m seawards of the land-water interface. This was generated by calculating mean baseline angle over a 200 m interval, with the transect location at the midpoint. RMSE (Equation (3.8)) was used to measure the distance between lines and MAE (Equation (3.9)) was used to determine net landward (positive) or seaward (negative) bias in prediction contours. MAE values were assigned as negative if the predicted contours were consistently seaward of the line derived from ground-reference measurements or manual digitisation.

The pixel-based evaluation metrics used were U_A (Equation (3.3)), P_A (Equation (3.4)) and F_1 (Equation (3.5)). All three metrics are suited to classification tasks with imbalance in class populations (e.g. non-edge pixels constitute more than 90% of the image). U_A values are more sensitive to the detection of inland non-coastal boundaries, so are typically lower than P_A values. A pixel incorrectly predicted to be the vegetation line will be classified as a false positive pixel, irrespective of the distance from the manually digitised or ground referenced line. To account for 'near-misses', where HED predicts the vegetation line to be at pixels close to the ground-referenced or manual digitisation measurements, the manually digitised and ground referenced lines were buffered to be three pixels wide (instead of one). Relaxed user accuracy, producer accuracy and F_1 scores were calculated by comparing HED outputs to the buffered ground referenced and manually digitised vegetation line measurements.

4.2.3.1 Validation image locations

Seven sites were used for HED validation (Table 4.1). High resolution ground measurements were collected from three of these seven locations along the Suffolk coastline of eastern England on 7 September 2019 (Walberswick, Dunwich and Covehithe) using an RTK-GPS with horizontal positional accuracy of 30 mm. Soft sandy cliffs are located at Covehithe with sharp cliff-top edge vegetation lines (Brooks and Spencer, 2010). In contrast, a more complex vegetation line on a mixed sand and shingle barrier is present at Walberswick and Dunwich (Pye and Blott, 2006). To ensure at least one ground-referenced measurement per pixel, points were captured approximately every 2 m alongshore and whenever there was a notable change in vegetation line direction. At Dunwich and Covehithe, isolated vegetation patches situated in front of the continuous vegetation line were not demarcated. At Walberswick, two vegetation lines were generated from ground-referenced measurements: i) a landward continuous vegetation line and ii) locations of isolated seaward vegetation patches. Confidence contours were compared to both vegetation lines derived from ground-referenced measurements at this site.

Ground-referenced measurements were compared to HED vegetation line predictions generated from a 3 m resolution PlanetScope image, using distance and pixel-based evaluation metrics outlined above. The PlanetScope image was captured on 12 September 2019 and was previously unseen by the HED model. Between 7 and 12 September 2019 waves approached from a dominant north easterly direction and rarely exceeded 1 m significant wave height (maximum peak significant wave height at Southwold Approach was 1.45 m (Cefas, 2020)). Due to these wave conditions, there is a high degree of confidence that the vegetation line remained stable over this time-period.

The trained HED model was also used to predict the vegetation line position at four additional locations where ground-referenced measurements were not collected (Table 4.1). At these locations HED output prediction contours were compared solely to manually digitised vegetation lines. Images from two locations (Winterton, UK and Perranuthnoe, UK) were used during HED training but different image dates were used (training image dates: 2018 and 2019, testing image dates: 2010 and 2015). The other two locations were previously unseen by the neural network: Wijk-aan-Zee, The Netherlands, and a section of Bribie Island, the smallest and most northerly of three major barrier islands, northern

Moreton Bay, Queensland, Australia.

Table 4.1: Locations of Holistically-Nested Edge Detection validation images. Other columns provide information on dominant shoreline direction, spring and neap tidal ranges, dominant sediment type, geomorphology and climate at each site as well as whether ground-referenced measurements of the coastal vegetation edge were collected.

Location	Country	Tidal range (m)	Dominant vegetation type	Geomorphology	Climate	Ground-referenced data collected (Yes/ No)
Walberswick	UK (Suffolk)	2.5(spring) 0.5 (neap)	Psammosere dune vegetation	Dune	Temperate	Yes
Covehithe	UK (Suffolk)	2.5 (spring) 0.5 (neap)	Cliff top grasses/ agricultural crops.	Soft sandy cliffs	Temperate	Yes
Dunwich	UK (Suffolk)	2.5 (spring) 0.5 (neap)	Psammosere dune vegetation	Shingle dune	Temperate	Yes
Winterton	UK (Suffolk)	2.5 (spring) 0.5 (neap)	Psammosere dune vegetation	Dune	Temperate	No
Perranuthnoe	UK (Cornwall)	6.0 (spring) 2.0 (neap)	Cliff top grasses and Psammosere dune vegetation	Beach dunes and rocky cliff	Temperate	No
Wilk-Ann-See	The Netherlands	4.0 (spring) 2.0 (neap)	Psammosere dune vegetation	Dune	Temperate	No
Bribie Island	Queensland, Australia	2.0 (spring) 0.5 (neap)	Eucalyptus forest	Barrier island	Sub-tropical	No

4.2.4. Determining the optimum spectral band combination

The default VGGNet-16 weights can only be initialised in a HED model which accepts images with three spectral bands. The performance of the HED model was therefore independently trained using four different combinations of three spectral bands: RGB, RG-NIR, BG-NIR and GB-NDVI. Output predictions from the four HED models were compared to select the most appropriate model for vegetation line detection. Figure 4.5, 4.6

& 4.7 provide a comparison of HED performance using different spectral band combinations at three locations not contained in either the training or the validation dataset: Cromer, UK; Varela, Guinea-Bissau and Wyk auf Föhr, North Frisian Islands, Germany. The HED models trained on spectral band combinations RG-NDVI and RGB predicted every pixel in the image to be the coastal vegetation edge and therefore these models were rejected. Only the HED models trained on images with spectral band combinations RG-NIR and BG-NIR were able to discard pixels not pertaining to the coastal vegetation edge. The HED model trained on BG-NIR spectral band images was still unable to discard many non-edge pixels and as a result produced very low user accuracy results of 0.06, 0.02 and 0.02 at Cromer, Varela and Wyk auf Föhr respectively. In contrast, the HED model trained using spectral bands RG-NIR was able to predict the location of the coastal vegetation edge with a user accuracy of 0.26, 0.59 and 0.25 at Cromer, Varela and Wyk auf Föhr respectively. The HED model trained using images with RG-NIR spectral bands was thus used to form the basis of the VEdge_Detector tool.

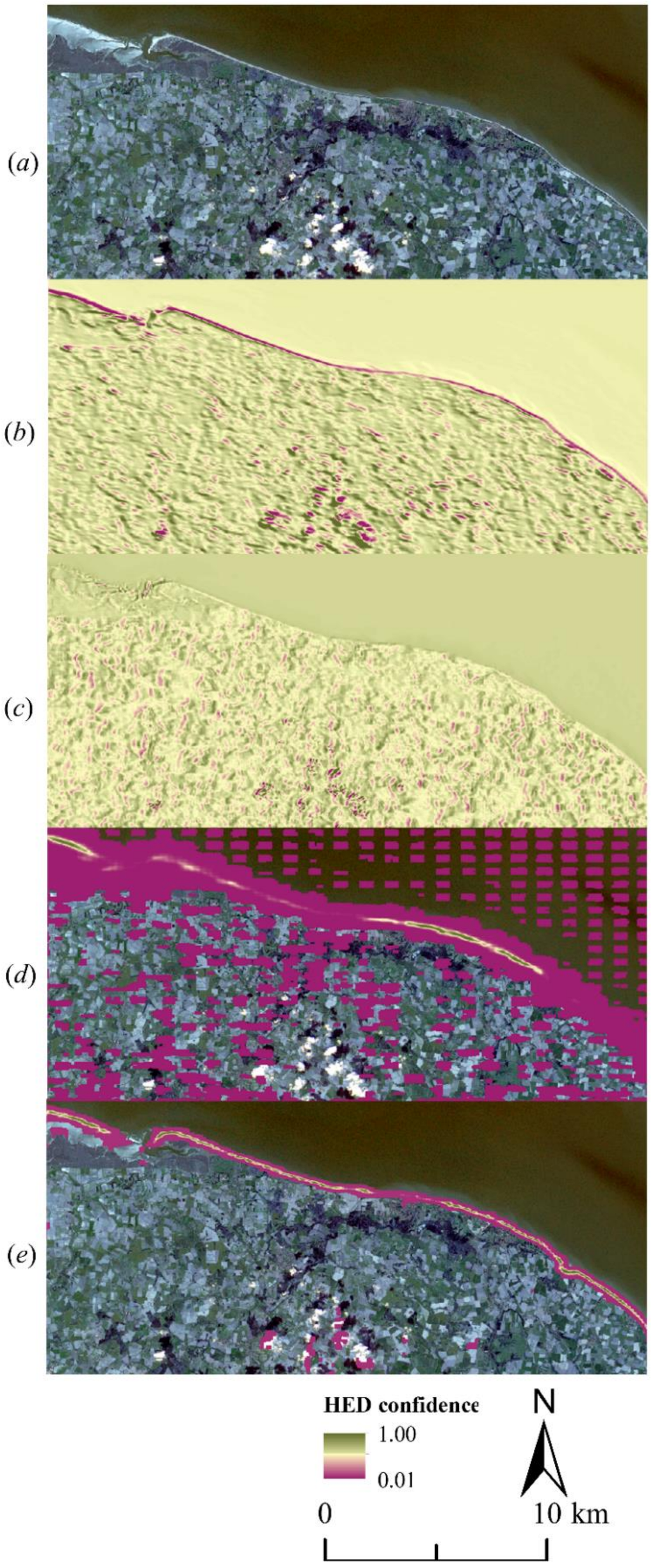


Figure 4.5: CNN predictions when trained with different spectral band combinations at Cromer, UK.(*a*) Original 3 m PlanetScope image of Cromer, Norfolk, UK (52°93’58.3 N, 1°27’18.0 E). Predicted coastal vegetation edge locations using the HED model trained with spectral band combination (*b*) RGB, (*c*) RG-NDVI, (*d*) BR-NIR, (*e*) RG-NIR.

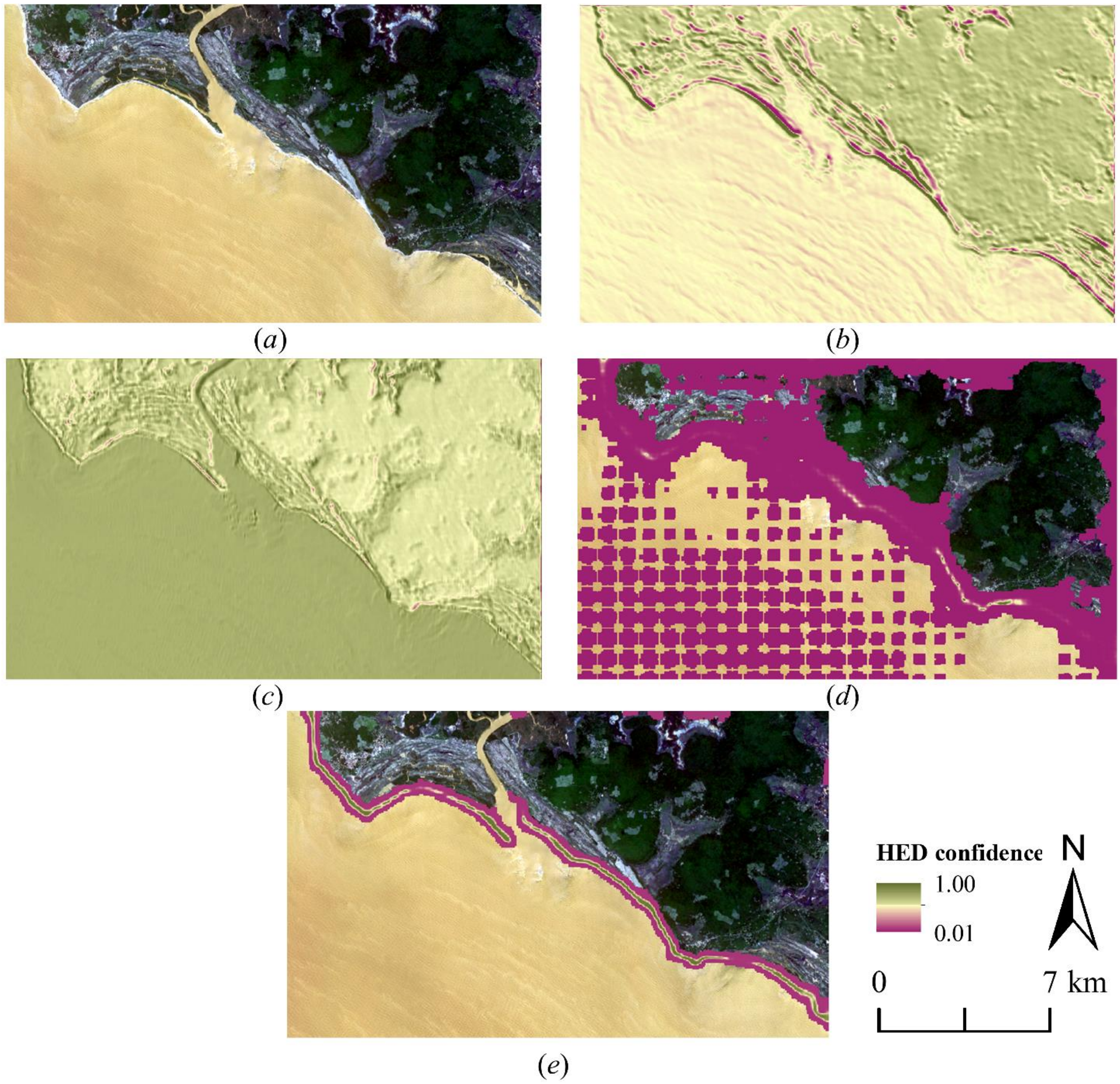


Figure 4.6: CNN predictions when trained with different spectral band combinations at Varela, Guinea-Bissau. (*a*) Original 5 m RapidEye image of Varela, Guinea-Bissau (12°28’61.0 N, -16°59’45.7 E). Predicted coastal vegetation edge locations using the HED model trained with spectral band combination (*b*) RGB, (*c*) RG-NDVI, (*d*) BR-NIR, (*e*) RG-NIR.

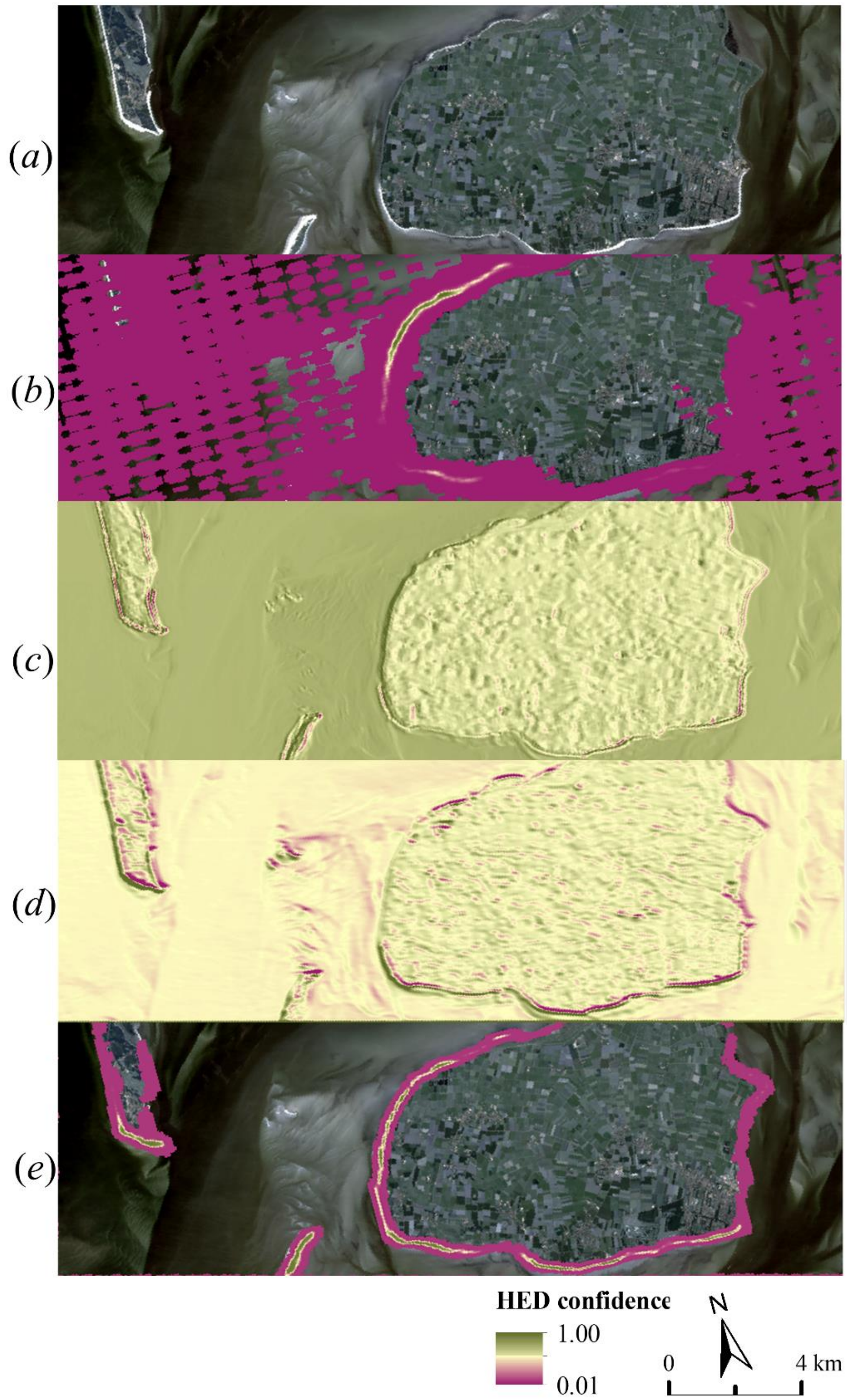


Figure 4.7: CNN predictions when trained with different spectral band combinations on the Frisian Islands, Germany. (*a*) Original 3 m PlanetScope image of the islands of Sylt, Amrum and Föhr, Frisian Islands, Germany (54°68’31.4 N, 8°55’74.4 E). Predicted coastal vegetation edge locations using the HED model trained with spectral band combination (*b*) RGB, (*c*) RG-NDVI, (*d*) BR-NIR, (*e*) RG-NIR.

4.2.5. Shoreline change detection

The VEdge_Detector tool was used to predict the vegetation line from 11 images of Covehithe spanning the period 2010 to 2020. To minimise the influence of seasonal changes to vegetation line location, all selected images were captured in the period between May and August of each year. Confidence contours were generated at 0.1 intervals from 0.05 to 0.95, creating a total of 10 landward and seaward contours per image.

Vegetation line change was calculated using DSAS in ArcGIS 10.5.1 (USGS, 2018). The position of the 10 confidence contours for every year was determined along transects running orthogonal to the dominant shoreline direction. Transects were separated by 10 m alongshore intervals. Change in the position of the landward and seaward 0.95, 0.55 and 0.05 confidence contours were calculated to determine rates of vegetation line change. Metrics calculated were Net Shoreline Change (NSC = distance between the oldest and most recent shoreline position) and End Point Rate (EPR = NSC divided by the time interval in years). To minimise geometric errors, ten tie-points were used to ensure consistent georegistration in the 11 images used in the shoreline change analysis. The locations of stable anthropogenic structures, including road junctions and building corners, were used as the tie points and were distributed evenly over the images.

Vertical aerial imagery of Covehithe, provided by the Environment Agency with 10 to 50 cm resolution, was manually digitised (Environment Agency, 2020a). NSC values derived using DSAS were compared when using vegetation lines produced by the VEdge_Detector tool and manual digitisation of aerial imagery. Due to aerial imagery availability, NSC values were compared across five baselines: 2010 to 2011, 2013 to 2014, 2015 to 2016, 2016 to 2017 and 2017 to 2018.

4.2.6. Comparing shoreline proxies

To determine any difference in the shoreline dynamics detectable using different shoreline proxies, the change in position of the waterline and vegetation line was compared at three transects alongthe Covehithe cliffs. The waterline was derived using the NDWI threshold contouring algorithm derived by Vos et al., (2019b). The change in position of the water and vegetation line was calculated using the DSAS method outlined in section 4.2.5.

4.3. Results

4.3.1. Manual, ground-referenced and VEdge_Detection measurements

Manually digitised vegetation lines were consistently located close to ground-referenced measurements (Root Mean Square Error (RMSE) was 1.72 m, 4.13 m and 2.28 m for Covehithe, Walberswick (landward) and Dunwich respectively). All sites exhibited a landward bias in manual digitisation, with Mean Absolute Error (MAE) of 0.82 m, 3.83 m and 1.83 m respectively. Across all sites, more than 93% of transects recorded an error ≤ 2 image pixels (6 m). At Walberswick, where ground-referenced measurements of two vegetation lines were collected, the manually digitised line was located closer to the landward continuous vegetation line than the seaward isolated vegetation patches (manual digitisation to seaward measurements RMSE = 16.72 m and MAE = 13.83 m). VEdge_Detector performance was therefore subsequently compared to the landward ground-referenced measurements at Walberswick.

The VEdge_Detector tool extracted continuous vegetation edges at all three field sites (Figure 4.8). For every site the VEdge_Detector 0.95 confidence contours were less than 5 m from ground-referenced vegetation line measurements (see Table 4.2 for summary of all RMSE and MAE values).

At Walberswick and Covehithe, all ground-referenced measurements were located between, or seawards of, the 0.95 confidence contours (Figure 4.8 (*a*) – (*b*)). The VEdge_Detector tool performed best at Covehithe with ground-referenced measurements located closest to the seaward 0.95 confidence contour (RMSE = 2.71 m, MAE = -0.02 m). A landward bias in the landward 0.95 contour (MAE = 7.98 m) and a seaward bias in the 0.95 seaward contour demonstrates that ground-referenced measurements at Covehithe were primarily located between the 0.95 confidence contours. Ground-referenced measurements were closest to the seaward 0.05 confidence contour at Walberswick (RMSE = 4.46 m, MAE = -1.11 m). Most ground-referenced measurements were situated between the seaward 0.95 (MAE = 4.31 m) and 0.05 confidence contours. The larger RMSE and MAE values for landward confidence contours compared to seaward contours shows a slight landward bias in VEdge_Detector outputs at Covehithe and Walberswick.

Table 4.2: VEdge_Detector accuracy at the three field sites determined by pixel and distance-based metrics from ground-referenced measurements. Shaded pixels in the mean absolute error column represent a landward (green) or seaward (blue) bias respectively in VEdge_Detector predictions. Darker colours represent a greater landward or seaward bias. Red boxes indicate the confidence contours with lowest RMSE and MAE per site. VEdge_Detector outputs are shown in Figure 4.8.

Site Location	U_A	P_A	F_1	U_A relaxed	P_A Relaxed	F_1 Relaxed	Confidence contour		RMSE (m)	MAE (m)
Covehithe	0.16	0.87	0.12	0.53	0.82	0.32	Seawards	0.05	5.39	-5.93
								0.95	2.71	-0.02
							Landwards	0.95	7.96	7.96
								0.05	9.82	9.82
Walberswick (seaward)	0.04	0.59	0.03	0.21	0.62	0.16	Seawards	0.05	8.32	7.08
								0.95	9.73	9.89
							Landwards	0.95	21.44	21.44
								0.05	26.04	26.04
Walberswick (landward)	0.11	0.84	0.11	0.50	0.79	0.31	Seawards	0.05	4.46	-1.11
								0.95	4.83	4.31
							Landwards	0.95	16.56	16.56
								0.05	21.59	21.59
Dunwich	0.07	0.85	0.06	0.21	0.81	0.17	Seawards	0.05	24.51	-24.29
								0.95	18.81	-18.69
							Landwards	0.95	2.37	-1.03
								0.05	5.98	5.21

The relatively high producer accuracy scores at these two sites (Covehithe = 0.87, Walberswick = 0.84) demonstrate that VEdge_Detector correctly detected a large proportion of vegetation line pixels derived from ground-referenced measurements. However, the lower user accuracy (Covehithe = 0.16 and Walberswick = 0.11) shows that a number of pixels inland of the field derived vegetation line pixels were also being detected by VEdge_Detector.

In contrast to the other two sites, VEdge_Detector predictions were primarily seawards of ground-referenced measurements at Dunwich (0.95 landwards confidence contour RMSE = 5.98 m, landward MAE = -5.21 m). The field line was located very close to the 0.05 confidence contour (RMSE = 2.37 m, MAE = 1.03 m). Producer accuracy values at Dunwich were consistent with the other two field sites, although a lower user accuracy was recorded (producer accuracy = 0.85, user accuracy = 0.07).

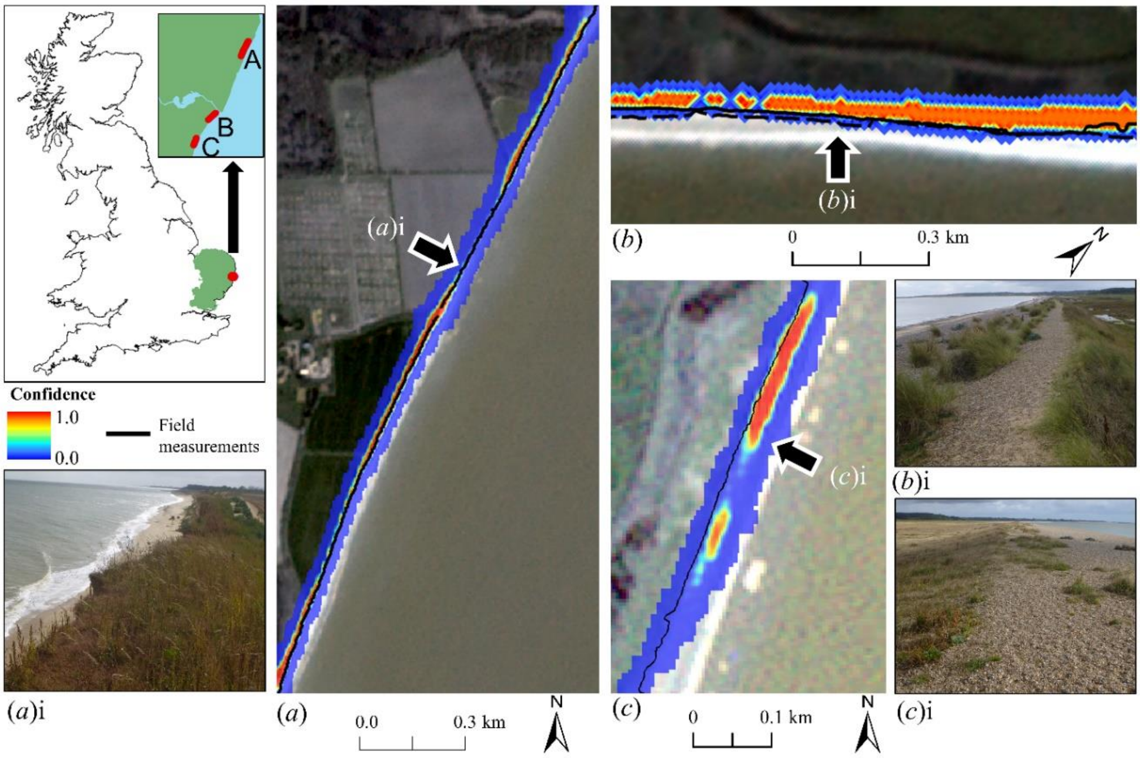


Figure 4.8: Comparison of VEdge_Detector tool predictions to field measurements of vegetation line at (*a*) Covehithe, (*b*) Walberswick and (*c*) Dunwich. Locations of photograph (*a*)i, (*b*)i and (*c*)i are show by arrows on corresponding images. The solid black lines show the field-delineated vegetation lines at all sites. At Walberswick, the landward and seaward vegetation lines derived from field measurements are denoted by a solid and dashed line respectively.

The VEdge_Detector tool produced a continuous vegetation line at three of the four sites without field data (Figure 4.9). The tool failed to predict a continuous vegetation line along some cliffed sections at Perranuthnoe, but a continuous line was generated along the beach sections and the cliffed sections to the right of the image (Figure 4.9 (*c*)). The tool performed best at Winterton and Bribie Island with errors less than 4 m between 0.95 confidence contours and manually digitised lines (Winterton MAE = -3.83 m, Bribie Island MAE = 3.11 m, Figure 4.9 (*a*) – (*b*), Table 4.3). Producer accuracy values greater than 0.9 were recorded at Winterton, Bribie Island and Wijk-aan-Zee, demonstrating a very high capability of the tool to detect the manually digitised vegetation line pixels. User accuracy was higher at Bribie (0.39) compared with Winterton (0.11), indicating that the tool produced a less precise line at Winterton.

User and producer accuracy values were lower at Wijk-aan-Zee and Perranuthnoe (Table 4.3) where more complex vegetation lines are present. More inland pixels were predicted as the vegetation line at these sites (Figure 4.9 (*c*) – (*d*)). There was a greater seaward bias in tool predictions at Perranuthnoe (RMSE = 7.14 m, MAE = -6.63 m) whereas distance-based error at Wijk-aan-Zee was comparable to Bribie and Winterton (RMSE = 4.61 m, MAE = 5.57 m).

Table 4.3: VEdge_Detector accuracy at the four validation sites without ground-referenced data determined by pixel and distance-based metrics. Shaded pixels in the mean absolute error column represents a landward (green) or seaward (blue) bias respectively in VEdge_Detector predictions. Darker colours represent a greater landward or seaward bias. Red boxes indicate the confidence contours with lowest RMSE and MAE per site. VEdge_Detector outputs are shown in Figure 4.9.

Site Location	U_A	P_A	F_1	U_A relaxed	P_A Relaxed	F_1 Relaxed	Confidence contour		RMSE (m)	MAE (m)
Bribie Island	0.39	0.98	0.28	0.41	0.99	0.29	Seawards	0.05	7.83	-6.35
								0.55	5.19	-3.55
								0.95	5.43	-3.01
							Landwards	0.95	3.72	3.11
								0.55	8.77	8.53
								0.05	12.3	12.11
Winterton	0.11	0.904	0.09	0.57	0.79	0.33	Seawards	0.05	32.37	-31.31
								0.55	21.59	-20.27
								0.95	7.97	-3.83
							Landwards	0.95	35.42	23.79
								0.55	50.39	48.94
								0.05	60.45	58.58
Wijk-ann-Zee	0.07	0.93	0.07	0.25	0.92	0.2	Seawards	0.05	14.56	-12.17
								0.55	8.87	-6.74
								0.95	5.26	-4.91
							Landwards	0.95	5.57	4.61
								0.55	8.66	6.86
								0.05	13.75	11.92
Perranuthnoe,	0.08	0.67	0.07	0.31	0.65	0.21	Seawards	0.05	36.12	-34.04
								0.55	35.55	-32.27
								0.95	7.14	-6.63
							Landwards	0.95	6.21	4.99
								0.55	9.98	8.69
								0.05	13.25	11.09

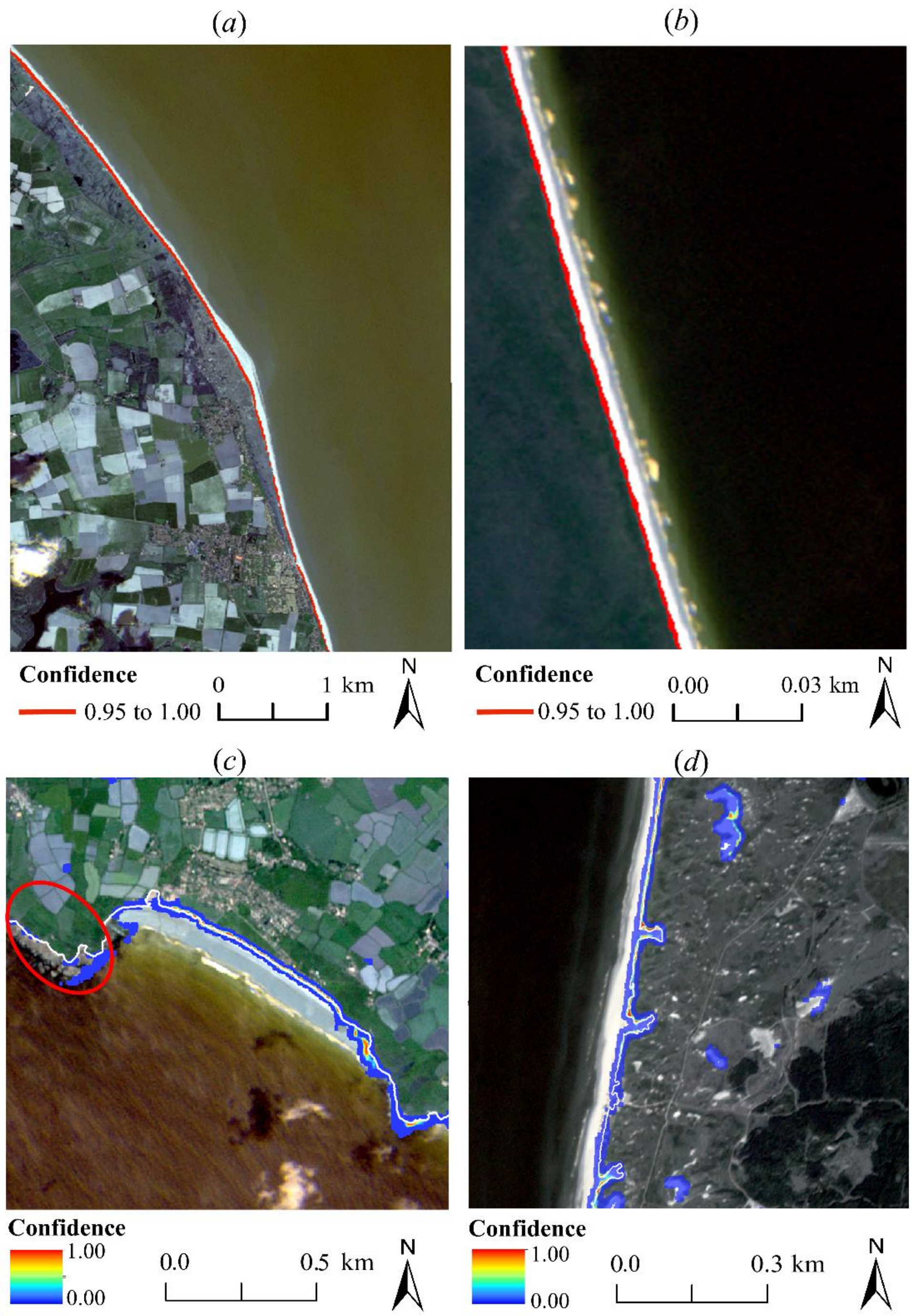


Figure 4.9. VEdge_Detector outputs at sites where ground-referenced measurements were not collected. (*a*) Winterton, Norfolk, UK (*b*) A stretch of Bribie Island, Australia, separate to the locations used for training outlined in the Supplemental Materials B (*c*) Perranuthnoe, Cornwall, UK. The red oval indicates the rocky cliff section where the VEdge_Detector failed to detect cliff top vegetation, (*d*) Wijk-aan-Zee, Netherlands. (*a*) and (*b*) display the predicted vegetation line in red with a confidence ≥0.95. (*c*) and (*d*) show examples of all VEdge_Detector outputs prior to applying any confidence thresholding.

4.3.2 Digital shoreline change analysis

For Covehithe, the VEdge_Detector tool generated confidence curves of vegetation line position from separate images captured in 2010 and 2020 (Figure 4.10 (*a*)). A continuous shoreline was extracted from both images, including where the vegetation line is interrupted by the local sand and gravel barriers that enclose Benacre Broad and Covehithe Broad. Total change in shoreline position between these two years was measured using the DSAS tool and the seaward 0.95 confidence contours (Figure 4.10 (*b*)). End Point Rates (EPR) along the Covehithe cliffs ranged between 2.47 m a^{-1} and 5.48 m a^{-1}, with an average retreat rate of 3.27 m a^{-1} (Figure 4.10(*b*)). The total amount of shoreline retreat during this period ranged between 24.27 m and 54.38 m; each transect with the smallest and largest retreat are shown as location A and B respectively in Figure 4.10 (*a*) – (*b*). Cross sections of the confidence curves at locations A and B are shown in the two insets in Figure 4.10 (*a*). The stretches of shoreline with the greatest rates of retreat corresponded to areas with no overlap in confidence curves. In contrast the confidence curves overlapped up to the 0.2 confidence contours at transects where retreat rates were lower.

The VEdge_Detector tool was subsequently used to generate confidence curves of vegetation line position at Covehithe annually between 2010 and 2020. Continuous vegetation lines were generated in all years except 2011, 2012 and 2018 when some agricultural fields had been ploughed, leading to apparent breaks in the vegetation line. The relative position of the annual confidence curves from 2010 to 2020 at the location with the fastest rate of retreat is presented in Figure 4.11. The vegetation line retreated landwards at a faster rate during the first half of the decade (End Point Rate (EPR) 2010 to 2015 = 6.92 m a^{-1}, 2016 to 2020 = 4.31 m a^{-1}, Figure 4.11). Individual years with the greatest rates of landward retreat were 2010 to 2011 (16.1 m ± 3.67 m), 2016 to 2017 (8.80 ± 3.24 m), 2013 to 2014 (6.93 ± 4.20 m) and 2017 to 2018 (5.31 ± 3.38 m). The smallest retreat rates were recorded in 2014 to 2015 (1.66 ± 2.45 m) and 2018 to 2019 (1.32 ± 3.44 m). The greatest distance between 0.95 landward and seaward confidence contours was in 2013 (6.70 m) and the shortest distance was in 2018 (1.23 m).

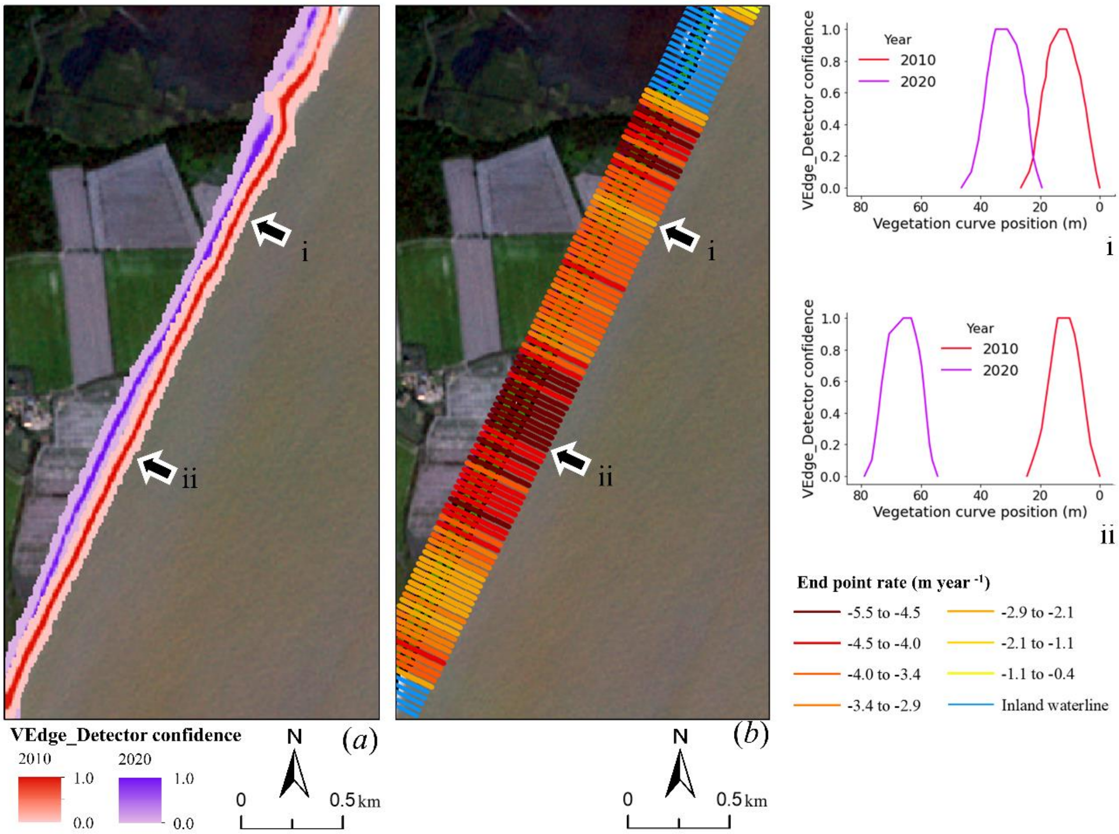


Figure 4.10: (a) VEdge_Detector outputs for a 2010 (red) and 2020 (purple) image of the Covehithe cliffs, Suffolk. Darker colours represent pixels predicted as the vegetation line with a higher confidence. Inset graphs, comparison of vegetation curves at transects situated at location i (smallest recorded change in shoreline position) and ii (largest recorded retreat in shoreline). Note: The image shows VEdge_Detector outputs with confidence values from 0.01 to 1.00, whereas the graphs show values 0.05 to 1.00 because the line graphs substantially 'fan' between 0.01 and 0.05. B) Rates of landward retreat (End Point Rate) at Covehithe between 2010 and 2020.

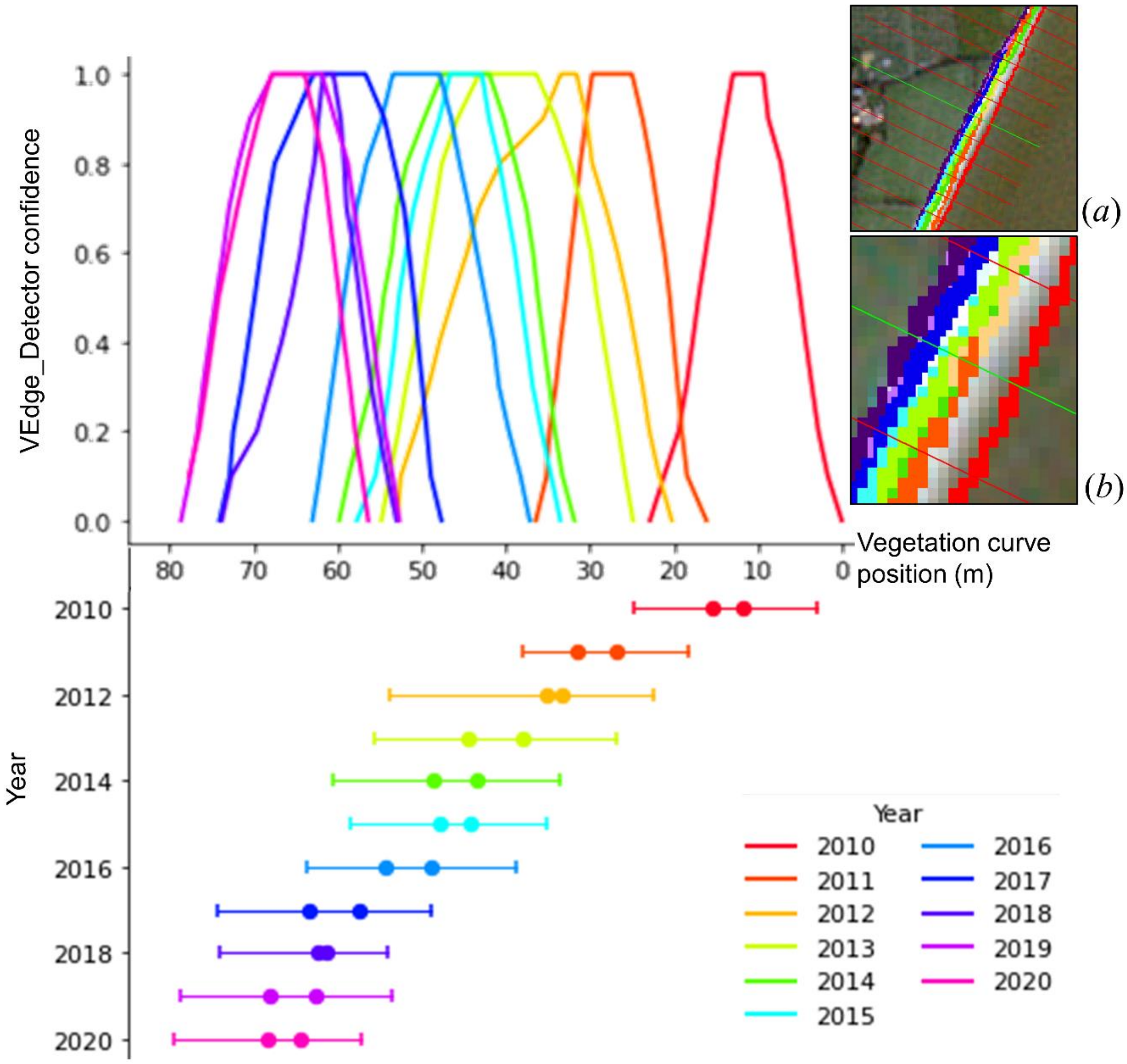


Figure 4.11: Shoreline change at Covehithe, Suffolk using VEdge_Detector outputs. Top: Vegetation confidence curve position during years 2010 to 2020 at one transect. Bottom: Representation of vegetation curves as a line. Dots represent locations of the 0.95 confidence contours, vertical lines represent locations of the 0.05 confidence contours. Insets i and ii: Transect location and all pixels predicted as the vegetation line with confidence greater than 0.95 overlaid on the 2020 image. Pixel colour coding by year is consistent with line graphs. Some of the colours are occluded in the image due to overlap.

Net Shoreline Change (NSC) values derived using DSAS were averaged across the entire

Covehithe coastline using both VEdge_Detector 0.95 confidence contours and manual digitisation of aerial imagery. Differences in NSC values obtained using the two methods ranged between 1.31 and 4.19 m, with a mean absolute difference of 2.19 m (Figure 4.12). An error value of ± 2.71 m was used for VEdge_Detector outputs, the RMSE between VEdge_Detector 0.95 confidence contours and ground-referenced measurements at Covehithe. Errors from digitising aerial imagery were set at 4.76% of each year's NSC value, consistent with calculations of error determined using the same digitisation method in Brooks and Spencer (2010).

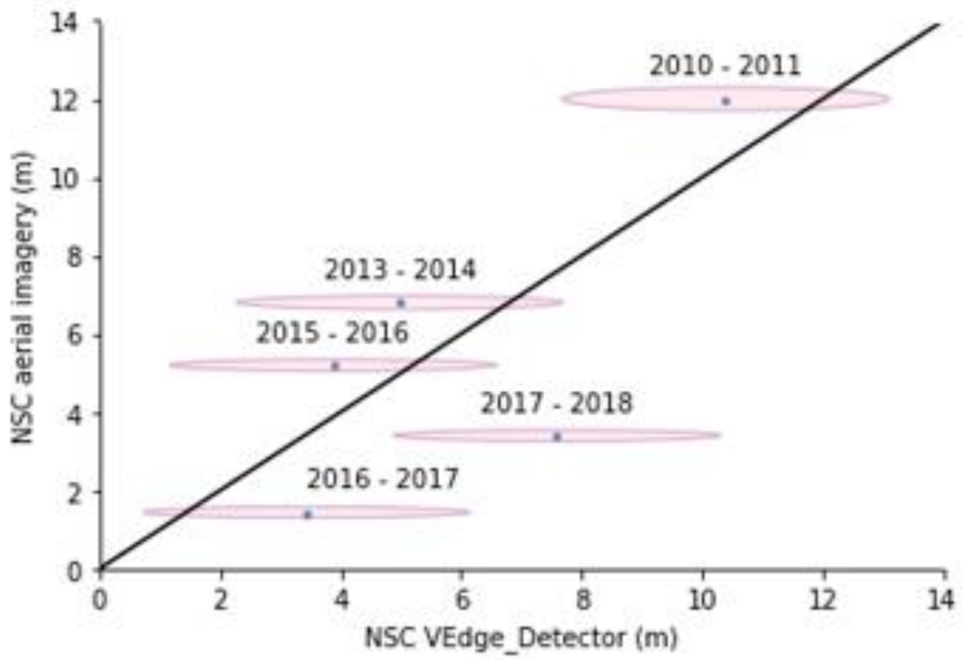


Figure 4.12: Comparison of Net Shoreline Change (NSC) values generated using VEdge_Detector 0.95 confidence contours and manually digitised aerial imagery. The blue dots show annual NSC values for the whole of the Covehithe coastline averaged over all orthogonal transects. The ovals represent the error associated with the two methods. The black line shows the position of the blue dots if there was an exact match between NSC values generated using the two methods.

4.3.3. Comparing shoreline proxies

The coastal vegetation edge consistently remained stable or retreated at every transect along the Covehithe coastline (Figure 4.13). In comparison, the waterline position fluctuated between periods of landwards and seawards migration. The waterline retreated along all transects between 2010 and 2014 but subsequently advanced seawards between 2014 and 2016. At all three transects, the vegetation line position in 2020 was landwards of its position in 2010. The waterline in 2020 was landwards of its position in 2010 at two of the three transects; however, at the site that had the smallest amount of retreat, the waterline in 2020 was seawards of its position in 2010 (Figure 4.13 (c)).

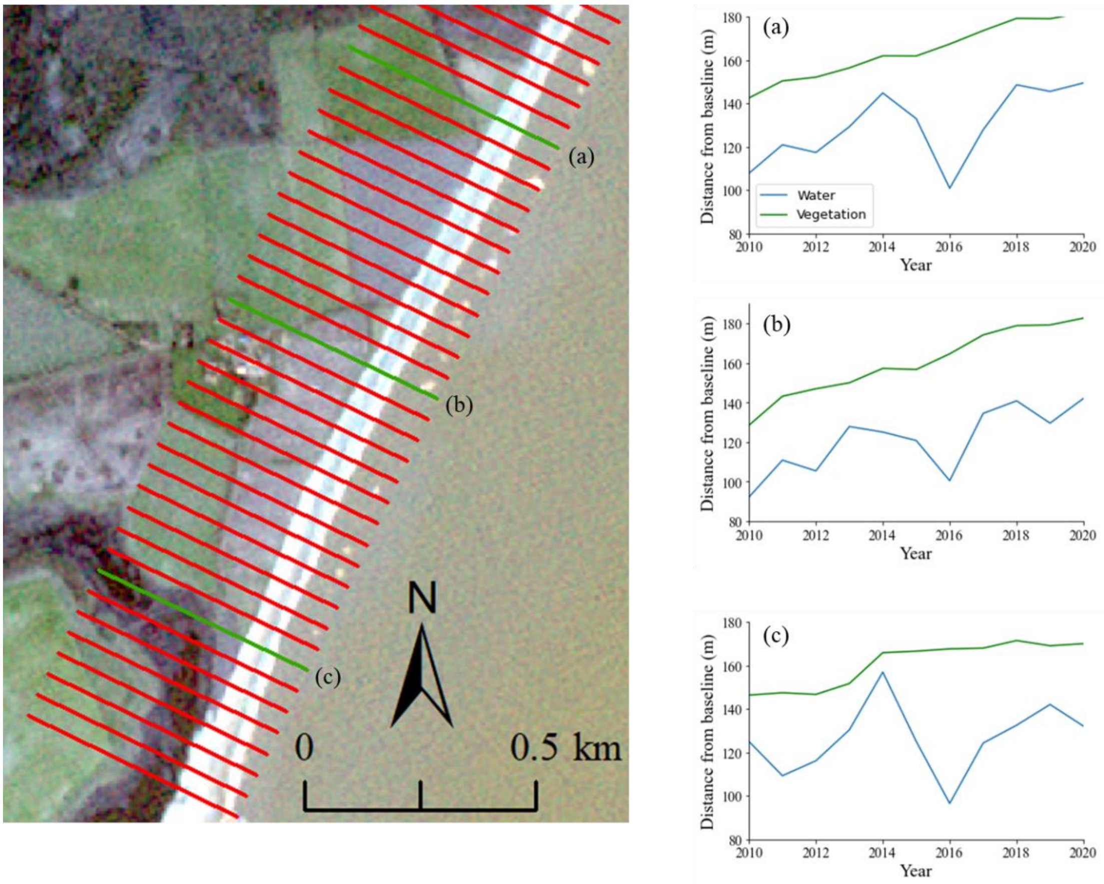


Figure 4.13: Comparison of the change in the position of the water and vegetation line at three transects, (a) – (c), across the Covehithe cliffs between 2010 and 2020. Transects overlaid on 3 m resolution PlanetLab image of Covehithe from 2016.

4.4. Discussion

4.4.1. VEdge_Detector performance

VEdge_Detector is the first fully automated tool for the digitisation of the coastal vegetation line from optical remote sensing imagery, where a trained Convolutional Neural Network (CNN) is used to detect the coastal vegetation line. The tool has been adapted from the Holistically-Nested Edge Detection (HED) model (Xie and Tu, 2015), a CNN trained to identify all objects in natural images. Here HED has been retrained to identify exclusively coastal vegetation edges, achieved by training the HED model on a comprehensive set of coastal remote sensing images. At six of the seven validation sites, VEdge_Detector 0.95 confidence contours were less than 6 m from coastal vegetation edges derived from ground-referenced measurements or manual digitisation of aerial imagery (Table 4.2 and 4.3). Previous studies have employed semi-automated methods to detect coastal vegetation, including thresholding and image classification (Zarillo et al., 2008; Rahman et al., 2010). VEdge_Detector advances these studies by being able to identify the coastal vegetation line in isolation, without requiring further post-processing steps to remove inland vegetation land covers and edges.

VEdge_Detector differs from other shoreline change studies by exclusively using Planet imagery with 3 m and 5 m spatial resolution. The combined high temporal and spatial resolution and coverage of Planet imagery provides a step-change in the ability to conduct shoreline change analysis. Previous studies have been primarily limited to digitising shoreline position in Google Earth Engine's Landsat or Copernicus imagery with 30 m and 10 m resolution respectively (Gorelick et al., 2017). Improvements in error values when using this imagery have been achieved using soft-classification, contouring and other methods with sub-pixel precision (Foody et al., 2005; Li and Gong, 2016; Pardo-Pascual et al., 2018). Extraction of the coastal vegetation line using imagery with 10 to 30 m resolution will remain problematic as one pixel can span the entire width of the coastal zone, incorporating numerous shoreline proxies. RMSE values derived in this study (2.37 m to 7.97 m) are comparable or a substantial improvement to error values derived from sub-pixel precision methods applied to coarser resolution imagery.

The combination of the high (up to daily) temporal resolution of the Planet imagery with the VEdge_Detector tool gives new opportunities to analyse the horizontal change in shoreline position caused by an individual major storm event or a succession of storm events (Roy, 2017). Previously this has only been possible through field or aerial based studies (e.g. Spencer et al., 2015). Studies of this nature are rare because data collection methods are time consuming, costly and information on shoreline position and profile prior to the storm event is only available in isolated, data-rich areas. The passive nature of image data collection used in VEdge_Detector, combined with its high spatio-temporal resolution opens new possibilities to assess storm damage, or other discrete erosion or accretion events in relatively understudied or inaccessible areas.

VEdge_Detector performed best on relatively simple, straight stretches of shoreline (e.g. Covehithe, Winterton and Bribie Island; Figure 4.8 (*a*) – (*b*)). Perranuthnoe, Cornwall, UK was the only location where VEdge_Detector did not generate a continuous vegetation line. This can be primarily attributed to the additional presence of rocky cliffs, because the majority of training data images contained only beaches. Whilst additional HED training using more images containing rocky cliffed shorelines may improve model performance, this may be at the expense of performance along sandy beached sections. Figure 4.9 (*c*) shows that the tool can detect a vegetation line at the base of some of the cliffs at Perranuthnoe, possibly due to the presence of macroalgae on the shore platform. It is beyond the scope of VEdge_Detector to include these sections, because change in macroalgal cover is highly unlikely to reflect an actual landward or seaward migration in shoreline position. Similarly, fixed coastal defences will not contain a mobile vegetation edge. Hence it is important to note that VEdge_Detector is primarily a tool for efficient and rapid extraction of the vegetation line from beach and dune systems over wide spatial coverage, from which shoreline change analysis can be performed.

The small discrepancies presented here between manually digitised shorelines and ground-referenced measurements from the first three cases studies provided confidence in using manually digitised shorelines to assess VEdge_Detector performance at several alternative sites where ground-referenced measurement was not possible (RMSE = 1.72 m, 4.13 m and 2.28 m for Covehithe, Walberswick (landwards) and Dunwich respectively). At Walberswick, the manually digitised line was closer to the landward fieldvegetation line measurements, indicating that manual digitisation primarily detects the more continuous vegetation line boundary landwards of the habitat of pioneer species. It appears that the diffuse nature of the vegetation edge in some locations, with isolated, dis-continuous vegetation clumps, can lead to discrepancies between manual digitisation, ground-referenced measurements and VEdge_Detector results because the best available imagery is of 3 to 5 m resolution.

This study has further shown that the method is robust at detecting the vegetation line on both tropical and temperate coasts. To date, the only previous use of CNNs to extract shoreline position was limited to a single location (H. Liu et al., 2019). Results presented here for seven different validation sites have shown that at six sites, producer accuracy was above 0.85, but user accuracy was lower than producer accuracy at every site (Table 4.2 and 4.3). This demonstrates how the tool is competent at correctly predicting the vegetation line pixel derived from ground referenced measurements but also generates a vegetation boundary region instead of a distinct line. These performance metrics are lower than those recorded by H. Liu et al. (2019) (user accuracy = 0.94, producer accuracy = 0.95). However H. Liu et al. (2019) used far coarser spatial resolution imagery (16 m to 50 m) and thus poorer user and producer accuracy results presented here could still result in lower RMSE values. Confidence contours were used throughout this study to determine where ground referenced measurements were located across predicted vegetation line confidence curves. At six of the seven sites, ground referenced measurements were closest to one of the 0.95 confidence contours, with RMSE less than 6 m (Figure 4.8 and 4.9). This highlights that even though a distinct vegetation line is not predicted, VEdge_Detector commonly predicts the ground referenced vegetation line with higher confidence than the surrounding pixels.

Vegetation lines were predicted with higher user accuracy along shorelines with abrupt vegetation edges. The fieldwork and additional validation sites with the highest user accuracy results were Covehithe (user accuracy = 0.16) and Bribie Island (user accuracy = 0.38) respectively. Bribie Island has an abrupt vegetation line as bare sand is found immediately adjacent to eucalyptus forest and Covehithe has an abrupt cliff-top vegetation boundary because cliff line retreat is too rapid for cliff toe vegetation establishment. In comparison, VEdge_Detector user accuracy results were lower at Dunwich (user accuracy = 0.07), Walberswick (user accuracy = 0.11) and Wijk-aan-Zee (user accuracy = 0.07) which all contain graded psammosere community vegetation on beach dune systems. The low user accuracy and higher producer accuracy results highlight how the vegetation edge is not a true line, but a boundary region graded from the presence of no vegetation to an increasingly dense vegetation cover when traversing inland. Discrepancies in the interpretation of vegetation line position occur even when reporting ground-referenced measurements. This was demonstrated at Walberswick where producer accuracy increased from 0.59, when using the most seaward pioneer vegetation, to 0.84 when using the landward continuous vegetation edge (Figure 4.8 (*b*), Table 4.2). Further investigation, supported with ground-referenced measurements, is required to determine whether this tool not only identifies vegetation edge location but also whether user accuracy results can indicate the degree of abrupt change in a vegetation boundary. An increase in vegetation line 'abruptness' can imply a loss of pioneer species seaward of dune systems, perhaps as a result of erosion under storm impacts or wave action associated with particularly high tides. Conversely, increasing widths in vegetation edge can represent relatively stable, or prograding, shoreline locations where vegetation has had the opportunity to establish and migrate seawards.

This Chapter also provides the first-ever comparison of the performance of a HED model using different spectral band combinations. RG-NIR visually outperformed other spectral band combinations, demonstrating the importance of spectral band selection in HED training. This finding is complementary to the universally applied vegetation detection algorithm, NDVI, which utilises the near infrared and red wavebands (Genovese et al., 2001). HED and many other CNN architectures only allow the input of images with three spectral bands (Simonyan and Zisserman, 2015). Improved performance may be achieved by concatenating the outputs of multiple CNNs trained on 3 band images. Marmanis et al. (2018) fused the outputs of two CNNs run in parallel, one CNN trained using spectral band information and the other trained using digital terrain models. Parallel CNNs were reported to automatically classify land covers with 84.8% pixel accuracy but no comparison to single CNN performance was provided. Further investigations should compare performance of single and multiple parallel CNNs trained exclusively on images with different spectral band combinations.

4.4.2. Shoreline change analysis using VEdge_Detector

The VEdge_Detector tool predicted a consistent landward shift in vegetation position between 2010 and 2020 at Covehithe, Suffolk (Figure 4.10). Years when the VEdge_Detector recorded the greatest rates of landward retreat coincide with North Sea storm surge events in December 2013 (Spencer et al., 2015; Wadey et al., 2015) and January 2017 (Floodlist, 2017) and the February to March 2018 ‘Beast from the East’ and ‘mini-Beast’ (Brooks and Spencer, 2019). Average rates of landward retreat at Covehithe derived from VEdge_Detector were consistent with results obtained in this study from manually digitising vertical aerial imagery. The mean difference in NSC values when using VEdge_Detector and digitising vertical aerial imagery was less than one pixel. These NSC values are also complementary to values derived along this stretch of shoreline using other proxy and datum-based methods (Brooks and Spencer, 2012; Burningham and French, 2017). This study has demonstrated the aptitude for the VEdge_Detector tool to accurately and efficiently detect the vegetation line from a relatively data rich shoreline where it has been possible to use other measurements, including aerial imagery, LiDAR data and Ordnance Survey data, to validate precision. Further applications of this tool should investigate its use in relatively data poor regions of the world or in regions where there is a necessity to determine the impact of coastal protection schemes or other anthropogenic interventions in the coastal zone.

A continuous vegetation line was generated at Covehithe for eight out of 11 years. During three years the vegetation line was fragmented due to the presence of ploughed agricultural land which interrupted the vegetation line. VEdge_Detector has been shown to be able to overcome issues of vegetation line fragmentation in other images, for example detecting a vegetation edge where agricultural fields had been harvested, although a fully continuous vegetation edge was not recorded every year at the landward extent of Benacre Broad and Covehithe Broad (Figure 10 (*a*)). Further studies should increase the ability of the tool to generalise, and use urban and water pixels when the vegetation line is fragmented. Alongside the sometimes fragmented nature of the vegetation line, it may remain an unsuitable proxy to use in shoreline change analysis in circumstances where there have been changes in vegetation communities as a result of both natural and anthropogenic processes unrelated to shoreline position. Therefore, this chapter suggests future research should combine multiple shoreline proxies simultaneously to provide a better indication of shoreline change.

4.5. Conclusions

This study has trained a Holistically-Nested Edge Detection (HED) model to produce VEdge_Detector, a fully automated tool for the extraction of coastal vegetation lines along sandy shorelines from optical remote sensing imagery. The semantic knowledge gained during HED training enables VEdge_Detector to discriminate between coastal vegetation edges and other inland vegetation boundaries, thus only extracting the coastal vegetation line and removing the need for subsequent post-processing. VEdge_Detector produces a vegetation confidence curve instead of a discrete line, which better represents how, in reality, the coastal vegetation line is not a distinct boundary but a broad zone where vegetation becomes a more continuous cover when traversing inland. The low error values (RMSE less than 6 m at all sites) between VEdge_Detector predictions and ground-referenced measurements demonstrates the aptitude for this tool to accurately detect the coastal vegetation edge location. VEdge_Detector performance varied depending on spectral band selection, with red, green and near-infrared shown to be the most pertinent image bands to use for coastal vegetation edge detection. This highlights the importance of image spectral band selection during CNN training in any context.

VEdge_Detector has been used to detect a decadal-scale, consistent landward shift in shoreline position at Covehithe, Suffolk, UK. This trend in vegetation line position is consistent with measurements obtained through manually digitising aerial imagery. This exemplifies how using this tool in different locations which exhibit a larger horizontal tidal range, may produce a more robust proxy of shoreline position than using the water line to determine net shoreline change. Despite the performance of VEdge_Detector, it has, to date, only been applied to relatively short (less than 100 km) stretches of coastline. Chapter 5 develops this work by investigating the performance of VEdge_Detector at a larger, supra-national scale.

This chapter is based on a published paper: Rogers, M.S., Bithell, M., Brooks, S.M. and Spencer, T. (2021). VEdge_Detector: automated coastal vegetation edge detection using a convolutional neural network. *International Journal of Remote Sensing*, *42*(13): 4805-4835. All analysis was my own work, except for the manual digitisation of the aerial imagery supplied by the Environment Agency, which was conducted by Dr Sue Brooks. This work is described in the last paragraph of Section 4.2.5 and the data contributed towards Figure 4.12

Chapter 5. Risk hotspots across the Guiana coastline, northern South America.

5.1. Introduction

Global-scale datasets pertaining to the distribution of different coastal risk elements are becoming increasingly available, including rates of shoreline change (Luijendijk., 2018; Mentashi et al., 2018), population densities (Stevens et al., 2015; Tatem et al., 2017), intertidal habitat vegetation distribution (Thomas et al., 2017; Hu et al., 2020) and land cover (Buchhorn et al., 2020). The interrogation and integration of these datasets can identify potential locations where humans are at greatest exposure or are at greatest risk to coastal hazards, in turn aiding the identification of locations to target more detailed study. The integration of global-scale shoreline change studies and population density maps, enables the identification of potential locations where the greatest number of humans are exposed to shoreline change (Figure 5.1). This integration highlights the Guiana coastline, northern South America as a location where a high proportion of the region's population is potentially exposed to shoreline change (Figure 5.1).

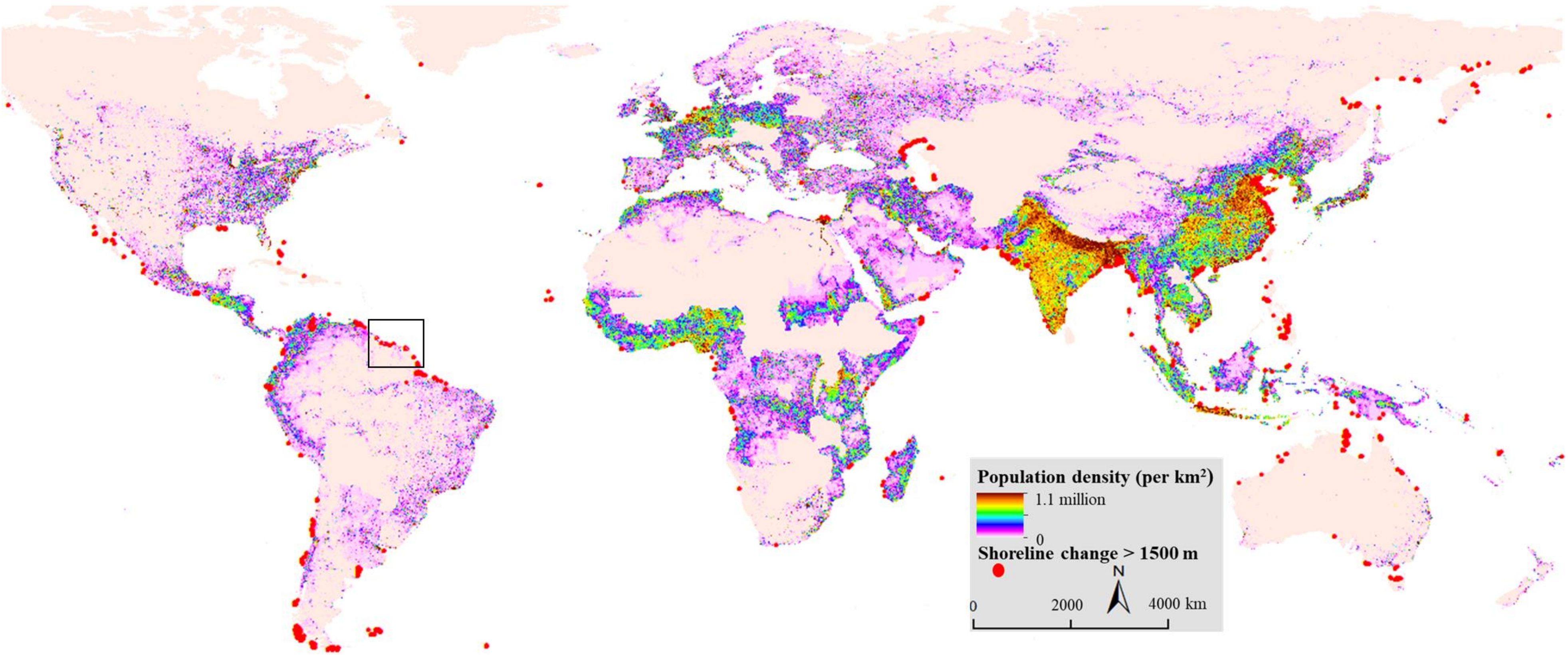


Figure 5.1: Global distributions of population densities and rates of shoreline change in the coastal zone. Colour ramp shows relative populations for every country in the world with a coastline, where brown, red and amber colours correspond to locations with high population densities, and blue, purple and white correspond to locations with low population densities. Locations with no values (peach background) correspond to locations with estimated population density < 1 person per km^2. Global population data was attained from WorldPop (Stevens et al., 2015; WorldPop, 2021). Red dots correspond to locations where shoreline change rates greater than ± 1500 m were recorded between 1984 and 2015. Shoreline change data adapted from Mentashi et al., (2018). Black rectangle corresponds to inset of the Guiana coastline in Figure 5.2.

The Guiana Coastline consists of the region between the Amazon and Orinoco Rivers. It includes the entire coastlines of Guyana, Suriname and the French overseas department of French Guyana (Hickey and Weis 2012; Figure 5.2). The Low Elevation Coastal Zone (LECZ) of South America, the hydrologically connected, contiguous area below 10 m above mean sea level, is projected to be home to 38 million people by 2060 (Neumann et al., 2015). In comparison, an estimated 180 - 220 million people currently reside in the LECZ of China, making the South American figure relatively small (Kulp and Strauss, 2019; Figure 5.1). However, alongside the total populations living in the LECZ, another important consideration is the proportion of each nation's population that live in the coastal zone.

Except for Belize, and island states such as the Maldives, Guyana and Suriname are the only countries in the world where 100% of their residents, living in urban zones, are situated in the LECZ (McGranahan et al., 2007; Colenbrander et al., 2019). Further, 90% of Guyana's population live below sea level (Vaughn, 2017), nearly the entirety of Guyana's agricultural production occurs within 25 km of the coastline (Hickey and Weis 2012), and more than 10% of the world's mangrove forest, important for mitigating incident wave energy, as well as providing medicine, fisheries, and fuel, is located in the region (Giri et al., 2010). These parameters highlight how acutely dependent the countries constituting the Guiana coastline are upon the coastal zone as a habitable space and a zone of economic activity. Identifying populations at the greatest risk to shoreline change in this region is, therefore, imperative for the targeted implementation of mitigation and adaption measures.

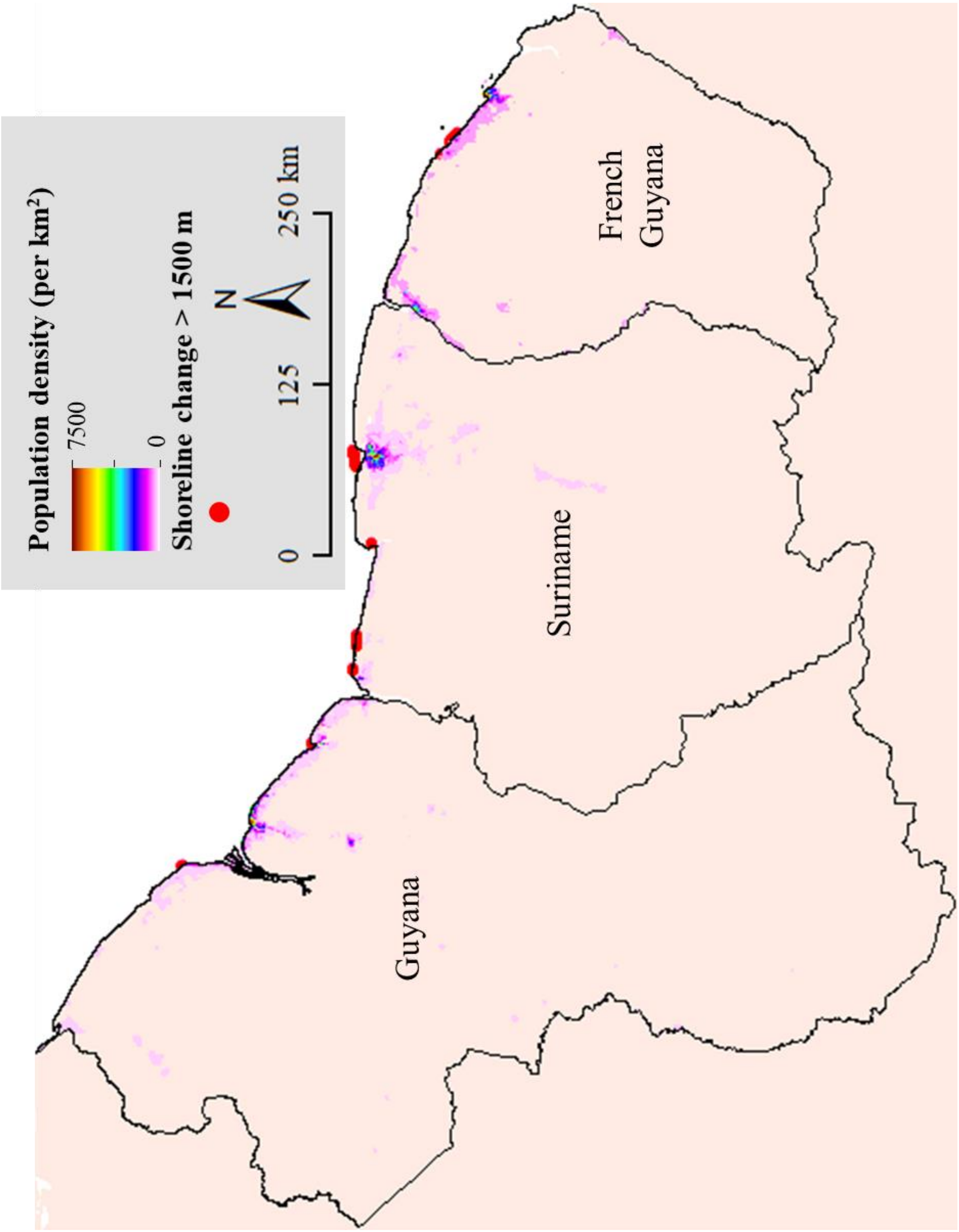


Figure 5.2: Population densities and locations with recorded rates of shoreline change greater than 1500 m between 1984 and 2015 across the Guiana coastline. Colour ramps and red dots correspond to the same features as in Figure 5.1.

The Guiana coastline is also one of the most dynamic shorelines in the world. It is characterised by the presence of mud banks measuring $10^{-1} – 10^{1}$ km in length, separated by 10 – 40 km wide inter-bank regions (Anthony et al., 2010; Walker et al., 2015; Spencer et al., 2016). Geochemical analysis has proven that the sediments accumulating on these mud banks originate from the Amazon River basin (Anthony et al., 2014), driven northwards by the trade wind-driven north equatorial recirculation region (Allison et al., 2000; Spencer et al., 2016). The spacing of the mud banks along the Guiana coastline has led to the hypothesis of a 30-year cycle in sediment deposition. The cycle begins with mud bank accretion and mangrove forest establishment followed by wave attack on the trailing edge of the mud bank, sediment erosion and mud bank migration (Allison et al., 2000; Allison and Lee, 2004; Plaziat and Augustinus, 2004; Spencer et al., 2016). At the wave exposed trailing edge of the mud banks, sediment is eroded, undermining the root system and eventually removing mangrove forest vegetation. Concurrently, sediment is deposited on the leeward, north westerly ends of the bank, resulting in alongshore mud bank migration at rates between less than 1 km (Froidefond et al., 1998) and greater than 5 km yr^{-1} (Gardel & Gratiot, 2005) and providing opportunities for renewed mangrove seedling establishment and forest development.

The 30-year cycle in mud bank extent has been attributed to three separate forcing mechanisms: the North Atlantic Oscillation (NAO), the El-Niño Southern Oscillation (ENSO), and the 18.6 year lunar nodal cycle (Chapter 1; Anthony et al., 2010; Gratiot et al., 2008; Walcker et al., 2015). A positive NAO and ENSO index is associated with mud bank recession across Guiana caused by high wave-energy events during the northern hemisphere winter (December – February) (Gratiot et al., 2008; Walcker et al., 2015). Conversely, mud bank accretion has been linked with years with a weakly positive or negative NAO or ENSO index, although no analysis of the influence of ENSO and NAO on mud bank extent since 2010 has been conducted. The 18.6 lunar nodal cycle causes Mean High Water Levels (MHWL) to vary by several centimetres (Walcker et al., 2015), which can also influence mud bank extent (Anthony et al., 2010). Further analysis is required to determine the relative importance of these extraneous factors in influencing mud bank extent along the Guiana coastline, and whether different factors are having a greater control in different localities.

Shoreline change along the Guiana coastline has been historically measured using ground-referenced surveys but extracting shoreline position from multispectral remote sensing imagery has recently become commonplace (Anthony et al., 2010). Studies using ground-referenced measurements conducted *in-situ* topographical mapping of mud bank morphology (Allison et al., 2000), used low flying airborne sensors (Lefebvre et al., 2004), and collected time-lapse ground-level photographs of mud bank position (Gardel et al., 2009). These methods collected high accuracy data on mud bank morphology and dynamics, but the inaccessibility and remoteness of many locations has precluded the use of these techniques at scales larger than $10^0 – 10^1$ km (Lefebvre et al., 2004). Recent increases in the temporal resolution of multispectral remote sensing imagery, combined with the availability of web-based platforms such as Google Earth Engine to analyse large volumes of satellite imagery (Gorelick et al., 2017), provides potential to map the dynamics of the entire 1500 km length of the Guiana coastline at annual to decadal timescales.

The most common proxy of Guiana shoreline position extracted from multispectral satellite imagery is the seaward extent of mangrove forests (de Jong et al., 2021; Gardel and Gratiot, 2005 & 2006; Jolivet et al., 2019; Walcker et al., 2015). Several mangrove species, including *Avicennia germinans*, are viviparous, enabling propagules released from adjacent forests to rapidly establish on, and stabilise, the new mud banks (Fromard et al., 1998 & 2004). This adaptive behaviour, combined with mangrove's ability to colonise the seaward limits of mud banks, makes the seaward coastal vegetation edge of mangrove forests a robust and commonly applied proxy of shoreline position along the Guiana coastline (Fromard et al., 2004; Gardel and Gratiot, 2005; Gratiot et al., 2008; Walcker et al., 2015). The instantaneous waterline position has also been estimated in French Guyana by classifying satellite imagery pixels into land and water classes (Bhargava et al., 2021). However, the automated extraction of the instantaneous waterline position along many stretches of the Guiana shoreline remains problematic because the shallow gradients of the mud banks make the waterline a difficult boundary to fix with any precision (de Vries et al., 2021).

To date, no study has mapped the seaward extent of mangrove forests across the entire Guiana coastline over multiple time periods. Fromard et al., (2004) monitored 50 years of mangrove extent and composition along one 20 – 30 km stretch of shoreline near Sinnamary, Suriname. The mangrove forests were shown to go through cyclical phases of erosion and accretion, supporting the theory of a thirty-year cycle in mud flat and mangrove extent (Fromard et al., 2004). These studies, whilst insightful, were limited to small areas of interest (less than 100 km), with no certainty that locations most vulnerable to shoreline change were identified. Further, supra-national scale studies of the Guiana coastline have the potential to identify dynamics that are not discernible in local scale studies, including the identification of spatially distinct stretches of coastline that exhibit similar shoreline change signals, and the tracking of the alongshore migrating nature of the mud banks.

Another important consideration is the use of repeated imagery to map shoreline change at a finer temporal-scale over the Guiana coastline. Mangrove position along the entire Guiana coastline has been calculated within global-scale studies of mangrove extent (Giri et al., 2011; Hu et al., 2020; Tang et al., 2018; Thomas et al., 2017; Bunting et al., 2018), with some studies highlighting the Guiana coastline as a hotspot for shoreline change (Thomas et al., 2017). However, all these studies only used imagery from one or two dates, preventing analysis of the change in rates of shoreline dynamics over multiple time periods. Detecting the Guiana shoreline position from satellite imagery captured during multiple years is required to identify locations that are consistently eroding or accreting, and areas that exhibit cycles of change in shoreline position.

To ascertain the differences in total risk levels that coastal communities are exposed to on this coastline, it is necessary to integrate metrics on shoreline change with data on population dynamics (Kron, 2013). Few studies have combined measurements of shoreline change and population density to identify risk hotspots in the Guianas. Proisy et al. (2021) developed a coastal vulnerability index (CVI) of the Suriname coastline as a function of the distance between the inhabited shoreline and the seaward mangrove extent, but no population metrics were contained in the CVI calculations. Population censuses for Guyana, Suriname and French Guyana have been conducted since 2012 (General Bureau of Statistics Suriname (GBSS), 2021; Government of Guyana, 2021; United Nations, 2021) but the datasets are heavily aggregated. Thus, for example, data from Guyana and Suriname is aggregated into just ten districts (GBSS, 2021; Government of Guyana, 2021). Finer resolution (100 m), grids of population density in the Guianas have been generated through dasymetric modelling (Stevens et al., 2015; Tatem et al., 2017). Dasymetric modelling uses Random Forests, a form of machine learning, trained to identify the relationship between population density and other ancillary parameters obtainable using remote sensing imagery and other datasets, including land cover, night lights and road density (Nagle et al., 2015). This modelling produces a finer, gridded representation of population density, validated using higher spatial resolution population data available from local authorities and other courses (Steven et al., 2015). Identifying changes in relative levels of humans exposed to shoreline change across the entire Guiana coastline has the potential to identify priority locations where resources might be directed towards coastal risk adaptation and mitigation measures.

This Chapter applies VEdge_Detector (Chapter 4), to automatically detect the coastal vegetation edge along the entire Guiana coastline from multispectral remote sensing imagery, developing the use of the tool from local to supra-national scale analysis (Rogers et al., 2021). The position of the seaward extent in mangrove vegetation was detected from Landsat imagery from five separate years between 1990 and 2020, Shoreline change rates were calculated and integrated with population density data (Stevens et al., 2015) to identify risk hotspots: locations with high population densities living near rapidly eroding shorelines. In three of these locations: Shell Beach, Guyana, Paramaribo, Suriname and Sinnamary, French Guyana, the vegetation edge between 2010 and 2020 was extracted from annual 3 – 5m spatial resolution Planet imagery. At these sites, the correlation between annual shoreline change rates and (i) NAO index; (ii) ENSO index and (iii) MHWL was calculated. Correlation analysis was conducted to locate stretches of spatially distinct shoreline that convey similar rates and direction of shoreline change to those identified at Sinnamary. Differences in the rates of shoreline change between each of the time periods was used to identify locations with shifting erosional hotspots.

5. 2. Methods

5.2.1. Study site and image selection

The Guiana coastline (northern South America) consists of Guyana, Suriname and French Guyana (Figure 5.3). This study used Tier 1 multispectral imagery from Landsat missions 5, 7 and 8, with 30 m spatial resolution, from which it is possible to detect shoreline change along the very dynamic Guiana Coastline (USGS, 2019). Landsat imagery was chosen for this project due to its relatively long (approximately 50 year) temporal coverage, from which it should be possible to observe the 30-year cycle in mudflat extent. Identification of cyclical shoreline change is not possible using Planet imagery because of its current temporal range of 10 years (Gorelick et al., 2017; Marta et al., 2018). Tier 1 Landsat data was used because it is consistently ortho-rectified and geo-registered, allowing imagery from the different Landsat missions to be compared (Young et al., 2017). The region spanned eight Landsat scenes (Figure 5.3) between coordinates (8.374 N, -59.85 E) and (4.011 N, -51.15E). Scene selection was conducted in Google Earth Engine (GEE) (Gorelick et al., 2017), and imagery was collected for the following five years: 1990 (Landsat 5), 1999 (Landsat 7), 2002 (Landsat 7), 2014 (Landsat 8) and 2020 (Landsat 8), resulting in a total of 40 Landsat scenes. Images from other years were not selected due to: i) a lack of concurrently captured imagery for all eight scenes (more than 3 month time difference between dates of scene capture); ii) high cloud cover; and iii) Landsat sensor failures (de Jong et al., 2021). Figure 5.4. provides a summary of the analysis conducted within this Chapter.

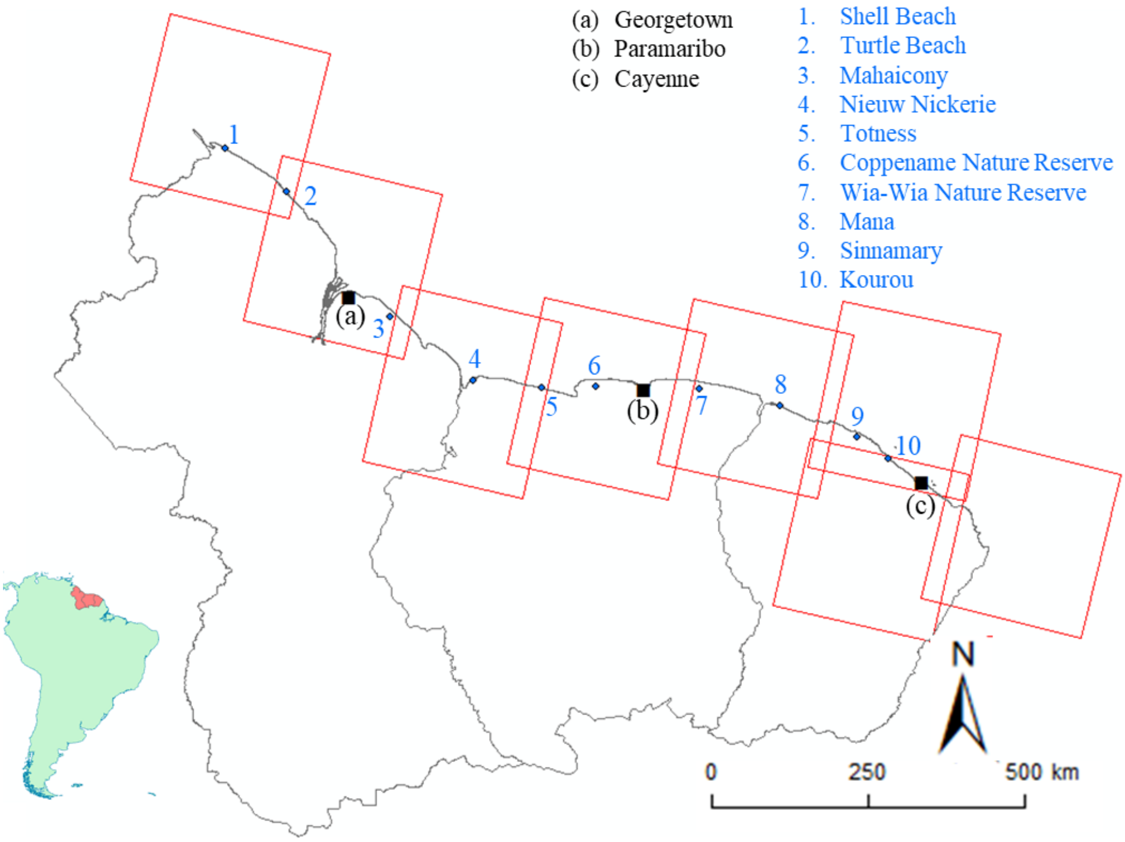


Figure 5.3: Study site of the Guiana coastline. Red squares depict the Landsat scenes used in this study. The capital cities of Guyana (Georgetown), Suriname (Paramaribo) and

French Guyana (Cayenne) are denoted by black squares. Blue points correspond to locations referred to in the chapter.

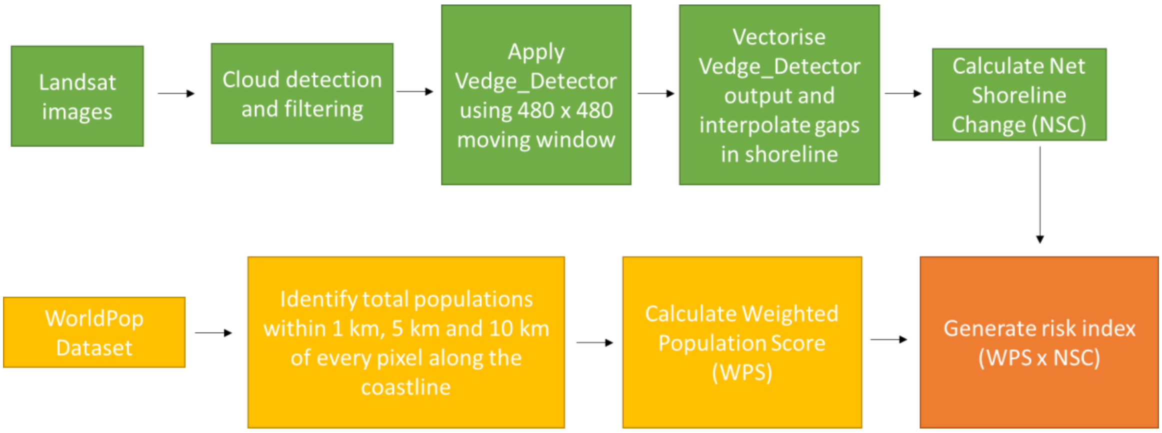


Figure 5.4. Overview of the steps taken to calculate rates of shoreline change, population densities and risk indices in the Guiana Coastal zone.

5.2.2. Image pre-processing: cloud detection and edge removal

Landsat images were filtered by percentage cloud cover in GEE. However, many images still contained clouds that needed to be removed to prevent their detection by VEdge_Detector (section 5.2.3). Clouds were initially detected in Landsat imagery using the *Landsat.simpleCloudScore()* function in GEE. This function detected clouds using Landsat's Thermal band and produced a binary cloud/non-cloud mask layer (Figure 5.5). The value of the pixels corresponding to clouds were modified to at-sensor radiance values typical of mangrove vegetation: green = 230, blue = 400, red = 150, near infrared = 3000. To smooth the cloud edges, a 15 × 15 kernel was convolved over the image, taking the average of the pixels in that kernel using Python's *Astropy* package. The pixels corresponding to clouds were assigned the values contained within the convolved image, with the non-cloud pixels being assigned the pixel values in the original image.

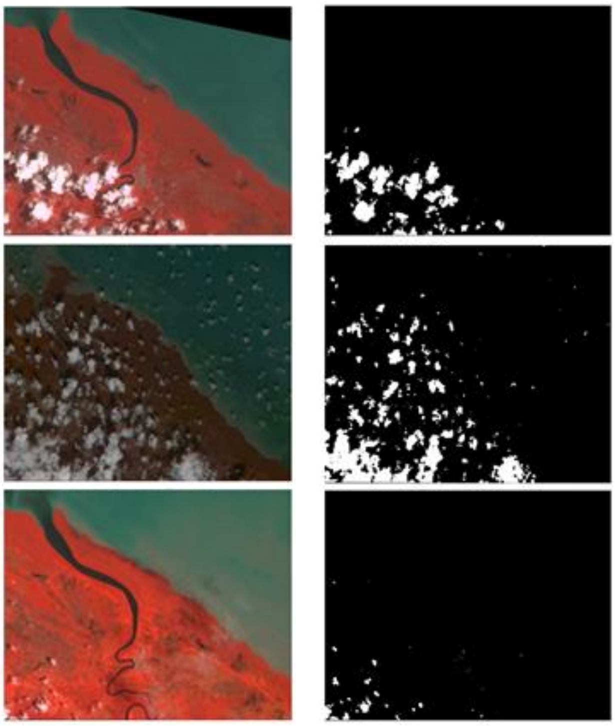


Figure 5.5: Cloud detection in Landsat imagery. Left: Original Landsat scene subsets of the Guiana coastline. Right: Corresponding binary cloud masks produced by the Landsat.simpleCloudScore() function in Google Earth Engine.

5.2.3. Shoreline detection across the Guiana coastline

5.2.3.1. VEdge_Detector

For every image, VEdge_Detector (Chapter 4) produced a raster layer with pixels valued between 0 and 1, representing the VEdge_Detector confidence that a pixel corresponded to the coastal vegetation edge. All VEdge_Detector predictions were performed on a Dell Laptop with Intel Core i7 and 32 GB of RAM. VEdge_Detector had a run time of approximately 15 minutes to detect the coastline in the Guiana study area, although VEdge_Detector training time far exceeds running time (Section 4.2.2).

5.2.3.2. Moving window algorithm

VEdge_Detector predicted a very wide coastal vegetation edge when applied to the original Landsat scenes and an additional pre-processing step was required to produce more precise predictions. The broad predictions were attributed to the Landsat scenes being far larger in size (greater than 30,000 km^2 or 7500 × 7500 pixels) than the imagery initially used to train VEdge_Detector (less than 10 km^2, 500 × 500 pixels). All images were resized to 480 × 480 pixels prior to VEdge_Detector input and it is argued that this downscaling is likely to have produced the very broad vegetation edge predictions.

To overcome these broad predictions, the shoreline was manually digitised in ArcGIS 10.5.1. A moving window algorithm was produced using Python's *Numpy*, *Gdal* and *rasterio* packages. This algorithm cropped the initial larger Landsat scene into multiple 480 × 480 pixel tiles, with the centre of a cropped tile positioned every 100 m along the manually digitised line. The use of an approximate, manually digitised line ensured that the cropped image was centred so as to include the coastline. The generation of an approximate shoreline was initially trialled using a well-established Normalised Difference Water Index (NDWI) threshold contouring (Chapter 3, Vos et al., 2019b). However, and particularly along stretches of coastline with shallow gradient muddy substrate, the NDWI threshold contour fluctuated erratically. As a result, not all images in the moving window algorithm were centred to contain the coastline (Figure 5.6). Some tile edge pixels were erroneously detected to be the coastal vegetation edge. Therefore, all pixels ≤ 10 pixels of the tile boundary were set to Not a Number (NaN). The tiles were subsequently mosaicked using Python's *Gdal.Warp()* function and the highest value was used for overlapping tile pixels.

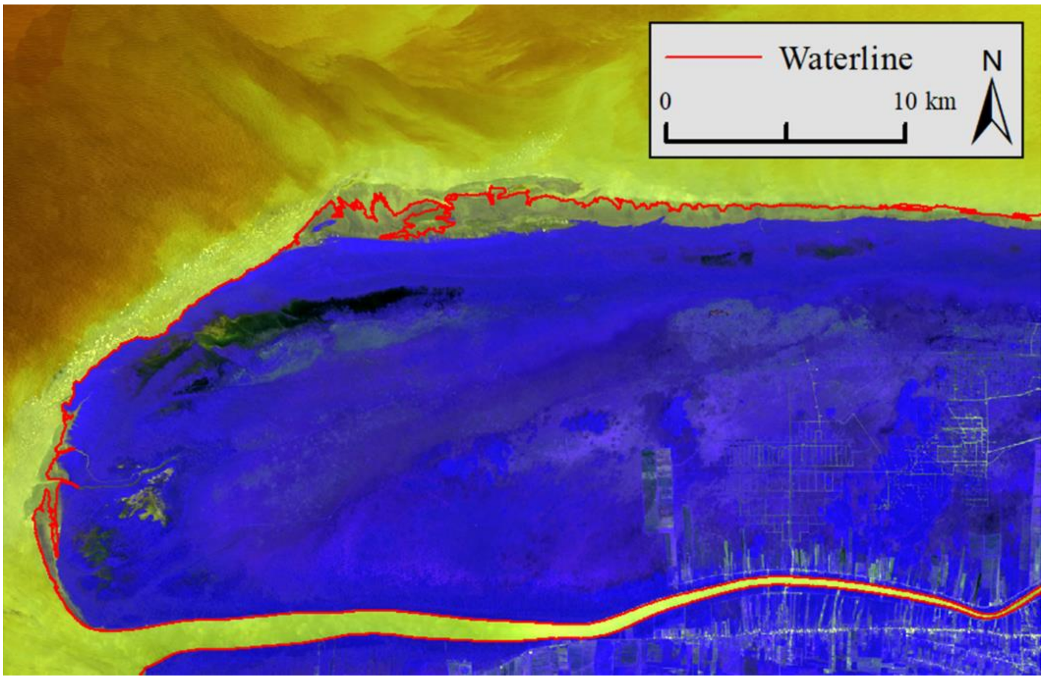


Figure 5.6: Waterline produced using NDWI threshold contouring method produced by Vos et al. (2019b) at Wia-Wia Nature reserve, Suriname. NDWI threshold contour overlaid on Landsat image captured in 1990 in red, green, near infrared false colour.

5.2.3.3. VEdge_Detector predictions post-processing

The shoreline change analysis method used in this study required the shoreline positions in vector polyline form (section 5.2.4). Python's *gdal.contourise()* function was used to convert the rasterised VEdge_Detector predictions into polylines. A continuous vegetation edge was detected along 90 - 95% of the Guiana Coastline for each of the five years in the analysis period 1990 – 2020. Gaps in the vegetation edge contours were interpolated by manual digitisation. To aid manual digitisation, the vegetation line was detected using the original red-green-blue image and the NDVI of the image. Tiles were inspected in ArcGIS 10.5.1 to identify and crop any contours corresponding to remaining clouds. The post-processed vegetation line contours were subsequently used in shoreline change analysis.

5.2.4. Shoreline change analysis using Landsat data

Shoreline change analysis was conducted using the ArcGIS plugin Digital Shoreline Analysis System (DSAS, Thieler et al., 2009; USGS, 2018). Net Shoreline Change (NSC = distance between the oldest and most recent shoreline position) and End Point Rate (NSC/ time in years) was calculated between 1990 and 2020 and for each of the intermediary time periods: 1990 – 1999, 1999 – 2002, 2002 – 2014, and 2014 – 2020. Change in the position of the shoreline was calculated along transects running orthogonal to the dominant shoreline direction. Transects were separated by 100 m intervals. NSC was used to compare rates of erosion and accretion over the entire study period. The End Point Rate (EPR) was used when comparing rates of change between the intermediary time periods. The length of the four time periods varied between three and 12 years, meaning EPR was a more suitable metric for comparing rates of shoreline change between the different time periods.

5.2.5. Weighted Population Score

To derive a population density value for each transect, 2020 population density maps at 100 m spatial resolution were sourced from the WorldPop dataset (Tatem et al., 2017). A moving window was passed over the input population images to calculate the total population living within 1 km, 5 km and 10 km of every grid square. A 10 km search radius was used because it is analogous to the width of the coastal strip containing most agricultural land and populations in the region (Hickey and Weis, 2012; Vaughn, 2017). To recognise that people living in closer proximity to transects experiencing shoreline change are likely to be more affected by the change, a Weighted Population Score (WPS) was calculated (Equation (5.1)):

$$WPS = (10 \times a) + (5 \times b) + (1 \times c) \quad (5.1)$$

where a, b and c correspond to the total population living within 1 km, 5 km and 10 km respectively of the grid cell. This analysis produced a 100 m resolution raster layer of WPS values for the entire Guiana region.

5.2.6. Identifying risk hotspots

To derive the WPS for each transect, the ArcGIS Sample function was used to find the value of the grid cell which was overlaid by each transect. To derive the risk index at each transect, the log normalised Weighted Population Score (WPS) and 1990 – 2020 Net Shoreline Change (NSC) values were multiplied together in the attribute table in ArcGIS (Equation (5.2)).

Datasets were initially log normalised to establish comparable ranges. The WPS was normalised to [0, 1] where zero and one represented locations with the sparsest and densest populations respectively. NSC values were normalised to [-1, 1], where positive and negative values corresponded to net accreting and eroding transects respectively and values near zero represented stable shorelines. To normalise the datasets, the log of all values was calculated and then all values were divided by the maximum log value to scale all values to between 0 and 1. For NSC, where positive and negative values existed, the log of the absolute values was calculated and all log values corresponding to eroding transects were multiplied by -1. Log normalisation ensured that the small number of transects with very large values did not mask the patterns in shoreline change and population density along the rest of the Guiana shoreline. Without this step, most transects had normalised values very close to zero.

To generate the Risk Index (RI) score along the Guiana shoreline, the log normalised WPS and NSC values were multiplied together (Equation 5.2). Despite the large number of studies investigating relative levels of coastal risk or exposure, there is no one method for integrating different risk indices (Rameri et al., 2011). This study adopted a similar method to that produced in the Coastal Risk Assessment Framework (CRAF) where hazard and exposure values are normalised and then multiplied together (Viavattene et al., 2018). This method was applied due to the even weighting it gives to the hazard and exposure metrics.

$$RI = NSC \times WPS \quad (5.2)$$

5.2.7. High resolution shoreline change analysis in hotspot locations

5.2.7.1. Planet imagery

At three locations with high Risk Index (RI) scores: Shell Beach, Guyana; Paramaribo, Suriname and Sinnamary, French Guyana, annual 3 – 5 m spatial resolution Planet imagery was collected. At Sinnamary, an image was collected from every year between 2010 and 2021, except 2015 when no Planet imagery was available. At Shell Beach and Paramaribo, one image was collected from every year between 2012, when the first image was captured from these sites, and 2021. VEdge_Detector was subsequently used to detect shoreline position in the Planet images using the moving window algorithm outlined in Section 2.3.2. EPR rates were calculated using DSAS as outlined in Section 2.4.

5.2.7.2. Cloud removal in Planet imagery

Most pre-existing cloud detection algorithms take advantage of the high reflective properties of clouds in the short-wave near infrared (SWIR) wavebands or low temperatures in thermal bands (Zhu et al., 2015; Sun et al., 2017; Frantz et al., 2018). Current Planet imagery does not contain a SWIR or thermal layer, so this study devised a thresholding algorithm which considered that clouds reflect relatively high amounts of radiation in the red and NIR wavebands (Figure 5.7). Pixels in the top 10% of at-sensor radiance values in both the red and near infrared spectral bands were designated as cloud pixels. A binary cloud mask layer was produced. After the cloud had been identified, the same process outlined in Section 5.2.2. was used to smooth the edges of the clouds: i) the values of pixels pertaining to cloud covered regions were altered to values typical of mangrove forest in Planet imagery; ii) a 15 × 15 smoothing kernel was convolved over the altered image to smooth the value of pixels at cloud edges; and iii) cloud-pixels were assigned the values contained within the convolved image, with non-cloud pixels being assigned the values in the original image.

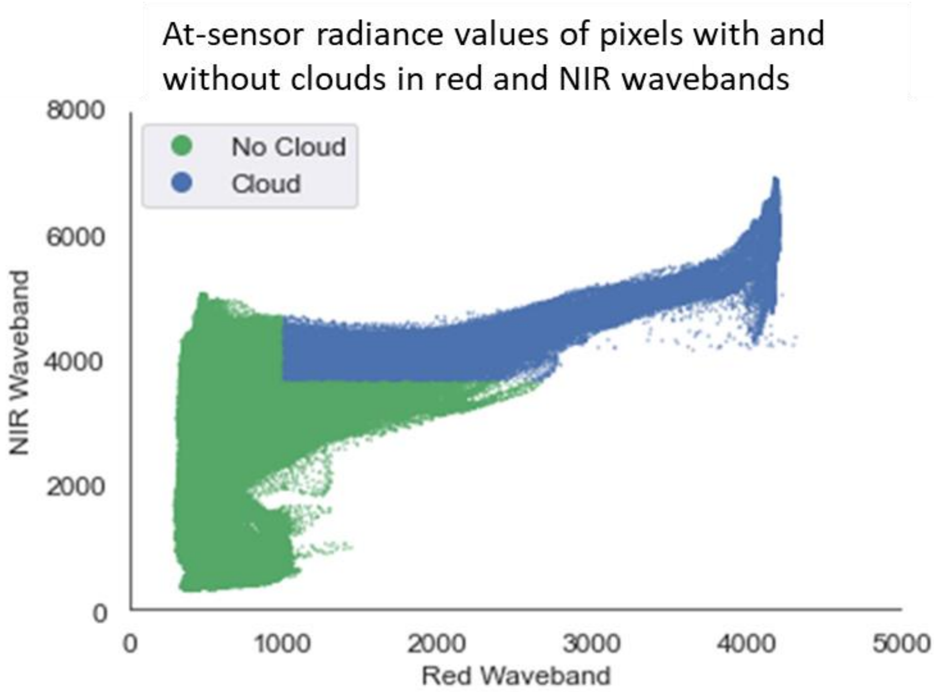


Figure 5.7: Comparison of the at-sensor radiance of pixels in red and NIR wavebands pertaining to locations with (blue) and without (green) clouds cover. The scatter graph contains the spectral property of approximately 4.5 million pixels.

5.2.8. Statistical analysis of shoreline change drivers

Correlations between 2010 – 2020 annual EPR values and i) NAO index; ii) ENSO index; and iii) MHWL at the Sinnamary and Paramaribo sites were calculated using Python's *Scipy* library. EPR values were statistically compared with the December January February March North Atlantic Oscillation (DJFM NAO) index, (University of East Anglia Climate Research Unit (UEA CRU, 2021), and the December January February El Niño Southern Oscillation (DJF ENSO) index (NOAA, 2021). The DJFM NAO index is calculated as the surface air pressure difference between Gibraltar and Iceland (UEA CRU, 2021) and the DJF ENSO index is calculated as the surface air pressure difference between Tahiti, French Polynesia and Darwin, Northern Territory, Australia (NOAA, 2021). These indices calculated the average index over the respective months and were chosen because the mud banks are most mobile during the northern hemisphere winter months (Anthony et al., 2014). Following Walcker et al. (2015), in order to account for lags in the shoreline response to conditions driven by the NAO/ ENSO index value, the average of the index score over the preceding five years was used. The statistical relationship between EPR and the extraneous forcing factors was investigated between 2010 and 2020 because this is the time period for which annual imagery of shoreline position was available.

To calculate the statistical relationship between EPR and the 18.6-year nodal cycle, annual MHWL were obtained from the Ile Royale tidal gauge positioned at coordinates (5.28 N, -52.58 E), approximately 20 km east of Sinnamary (National Oceanography Centre, 2021a). The Shapiro test for normality was applied to all datasets prior to conducting correlation tests (Dytham, 2011). The Pearson's correlation coefficient was used for normally distributed data, and Spearman's rank was applied to non-normally distributed datasets (Dytham, 2011). For all tests, the null hypothesis was rejected when $p < 0.05$.

5.2.9. Identifying regional scale dynamics

5.2.9.1. Correlations with Sinnamary site

Statistical analysis was conducted to determine whether other stretches of shoreline have exhibited the same response to the extraneous forcing factors as Sinnamary. The EPR values for all transects across the Sinnamary study site were averaged to provide an overall response of the mud bank to external forcing factors. The strength and direction of correlation between EPR values at Sinnamary and every other transect along the Guiana coastline was subsequently calculated. Correlations were calculated using the Python's *Pandas* library, using the Pearson's correlation test. Transects were considered to be correlated with the Sinnamary site values when the correlation coefficient value, r, was greater than 0.9 and the p-value < 0.05.

5.2.9.2. Erosion rates near Mana

To investigate the changes in erosion rates between 1990 and 2020 near Mana, French Guyana, EPR values at every transect for each of the four time periods were plotted on a line graph. The stretch of coastline around Mana contained approximately 700 transects and was greater than 70 km in length. The EPR values for every time period were plotted using Python's *Matplotlib* library.

5.3. Results

5.3.1. VEdge_Detector performance

VEdge_Detector successfully identified the coastal vegetation edge along more than 90% of the Guiana coastline during each of the five years for which Landsat imagery was available (1990, 1999, 2002, 2014 and 2020). For all years, a continuous shoreline was detected along the entireity of the coastlines of Guyana and French Guyana, with the exception of a small (c. 5 km) stretch of shoreline between Mana and Sinnamary, French Guyana. In Suriname, VEdge_Detector consistently failed to identify the coastal vegetation edge along two separate, approximately 30 km long, stretches of shoreline near the Coppename and Wia-Wia nature reserves.

5.3.2. NSC and EPR along the Guiana coastline

Net Shoreline Change (NSC) between 1990 and 2020 across the entire Guiana coastline ranged from 3200 m of accretion at the western edge of the Coppename Nature Reserve, Suriname to -2200 m of erosion near Sinnamary, French Guyana (Figure 5.8). In Guyana, the greatest rates of retreat over the 30-year period, -1700 m, were situated at Turtle Beach, and the most rapid accretion, 1140 m, occurred approximately 60 km south east of Turtle Beach. At Shell Beach, North West Guyana a stretch of accreting shoreline (NSC = 900 m) was located immediately to the west of a site of rapid erosion (NSC = -900 m) (Figure 5.9). Similarly in French Guyana, the very large rates of erosion (NSC = 1200 m) and accretion (NSC = -2200 m) occurred immediately adjacent to each other on the Sinnamary mud bank. The prograding shoreline was located at the western head of the Sinnamary mud bank, whereas erosion was predominantly found in the centre and western edges of the mud bank. The highest recorded rates of erosion in Suriname were found near Mana (NSC = -2100 m).

The proportion of stable shoreline, defined here as NSC less than ±60 m or two Landsat pixels over the 30-year period, varied substantially within each country. In French Guyana, 972 transects, or 21.9% of the country's shoreline experienced less than ±60 m net erosion or accretion. Two separate stretches of shoreline near the town of Kourou and the French Guyanese capital, Cayenne, were the most stable shorelines across the entire study area (NSC less than 15 m). In Guyana and Suriname, 9.3% and 14.8% respectively of the countries' shoreline were stable. In both countries, stable stretches of shoreline were predominantly clustered around river inlets and the Guyanese capital, Georgetown. Differences in NSC and EPR values between each county were calculated because the country boundaries correspond with major river systems, which produce natural breaks in the Guiana coastline, and correspond to marked changes in dominant shoreline direction.

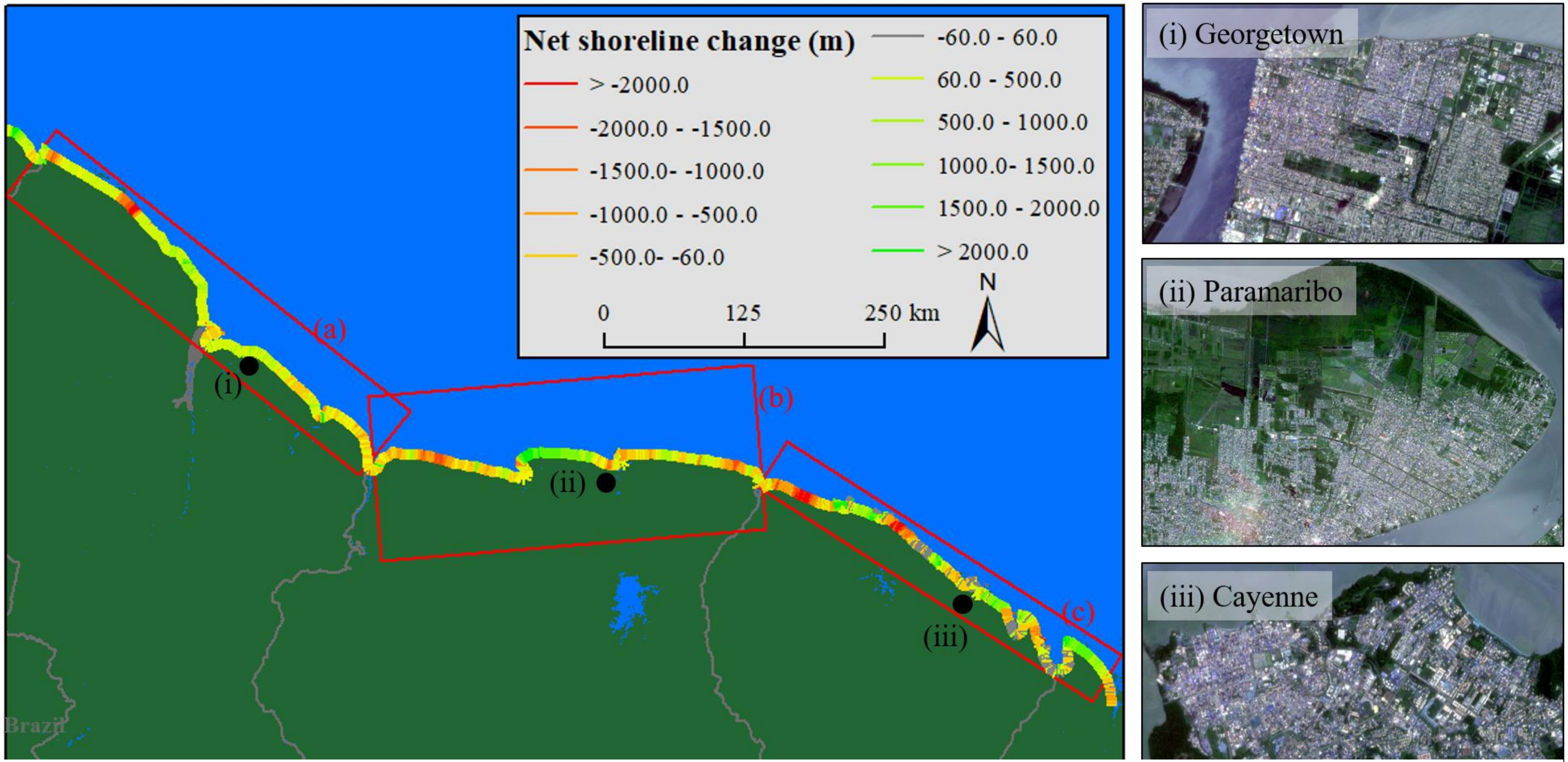


Figure 5.8: Net shoreline change rates across the Guiana coastline between 1990 and 2020. Red and amber transects correspond to sites of net erosion and green transects represent net accretion over this time period. Grey transects represent locations where NSC was < 60 m (2 Landsat pixels). Black dots (i), (ii) and (iii) show the locations of Georgetown, Paramaribo and Cayenne, shown in the figure insets. Red polygons denote the subsets presented at a large scale in Figures 5.8, 5.9 and 5.10. Shoreline change transects are overlaid on a 500 m resolution Terra MODIS surface reflectance image classified into land and water classes (ORNL DAAC, 2018).

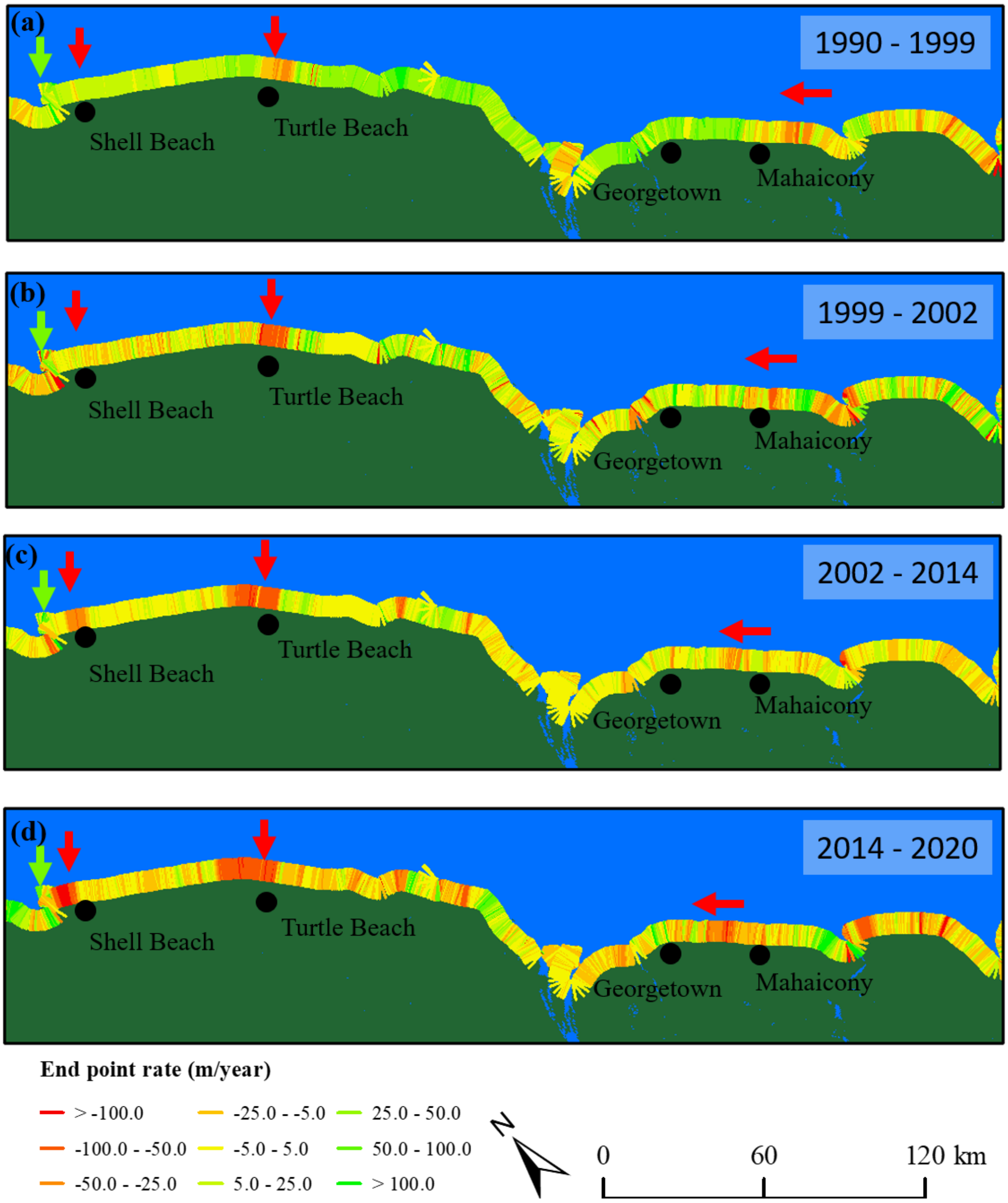


Figure 5.9: End point rates along the shoreline of Guyana (site (a) in Figure 5.7) between (a) 1990 – 1999, (b) 1999 – 2002, (c) 2002 – 2014 and (d) 2014 – 2020. Vertical red and green arrows highlight locations which experienced net erosion or accretion during all four time periods respectively. Horizontal red arrows highlight locations where a stretch of erosion migrated westwards in particular time periods between 1990 and 2020. Shoreline change transects are overlaid on a 500 m resolution Terra MODIS surface reflectance image classified into land and water classes (ORNL DAAC, 2018).

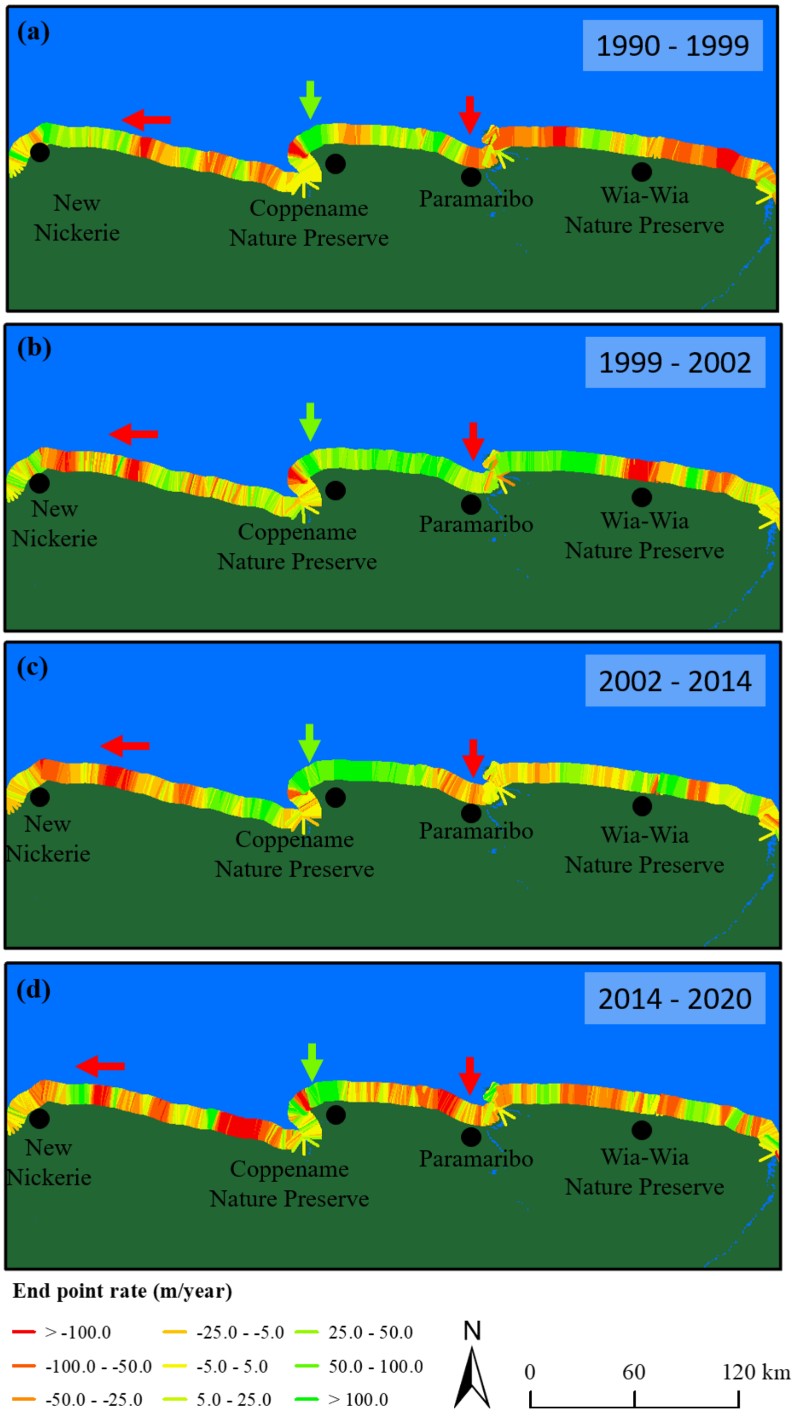


Figure 5.10: End point rates across the shoreline of Suriname (site (b) in Figure 5.7) between (a) 1990 – 1999, (b) 1999 – 2002, (c) 2002 – 2014 and (d) 2014 – 2020. Vertical red and green arrows highlight locations which experienced net erosion or accretion during all four time periods respectively. Horizontal red arrows highlight locations where a stretch of erosion migrated westwards in a particular time period between 1990 and 2020. Shoreline change transects are overlaid on a 500 m resolution Terra MODIS surface reflectance image classified into land and water classes (ORNL DAAC, 2018).

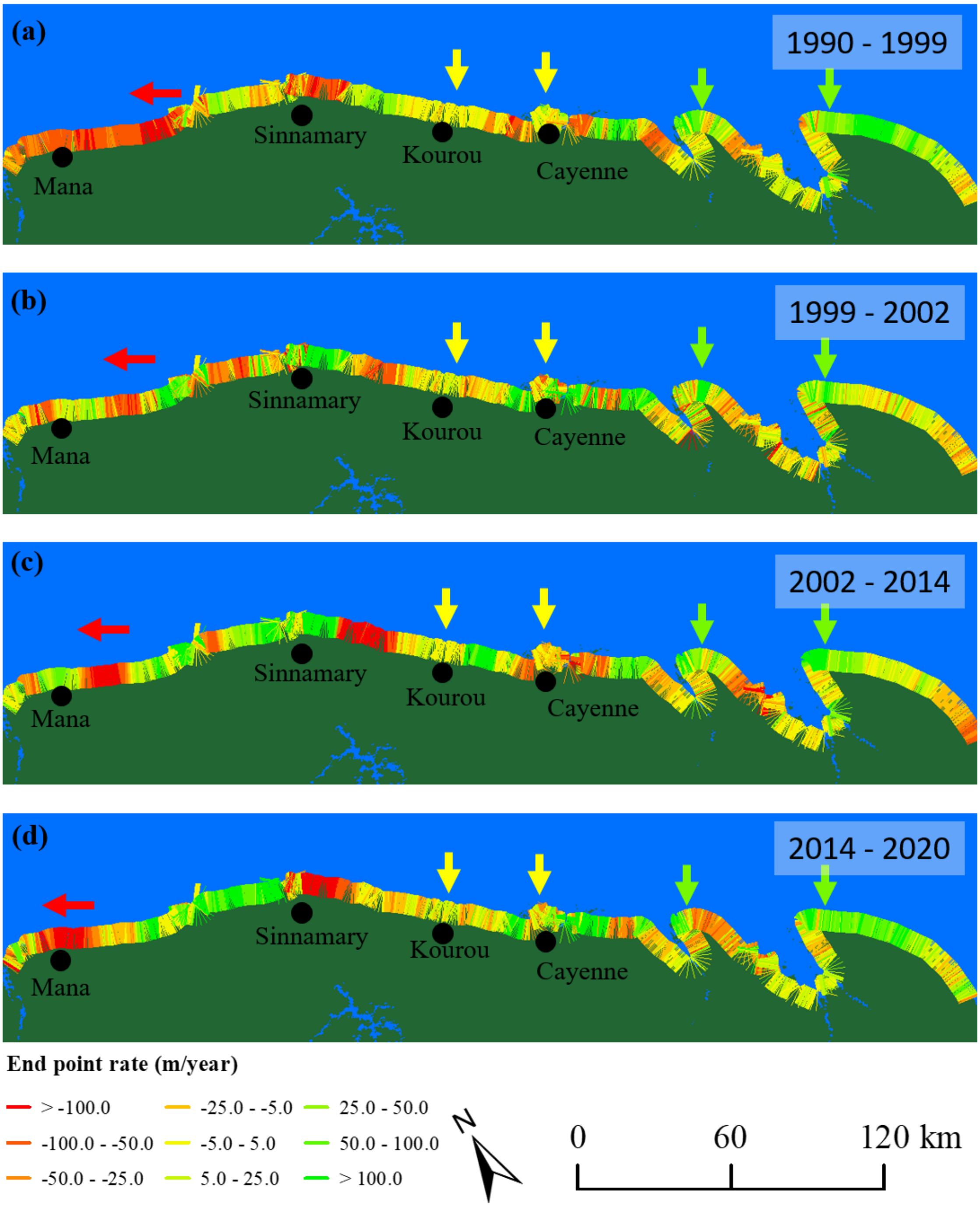


Figure 5.11: End point rates across the shoreline of French Guyana (site (c) in Figure 5.7) between (a) 1990 – 1999, (b) 1999 – 2002, (c) 2002 – 2014 and (d) 2014 – 2020. Vertical yellow arrows highlight shoreline locations which experienced negligible (less than 5 m year^{-1}) shoreline change between all time periods. Vertical green arrows highlight locations which experienced net accretion during all four time periods. Horizontal red arrows highlight locations where a stretch of erosion migrated westwards in a particular time period between 1990 and 2020. Shoreline change transects are overlaid on a 500 m

resolution Terra MODIS surface reflectance image classified into land and water classes (ORNL DAAC, 2018).

Many locations along the Guiana coastline experienced consistent erosion or accretion at each time period. Stretches of shoreline which accreted during every period included Western Shell Beach, Guyana (Figure 5.8) (EPR 1990 – 2020 = 20 - 26 m year^{-1}), Western Coppename Nature Reserve, Suriname (Figure 5.10) (EPR = 102 – 110 m year^{-1}) and the most easterly peninsula in French Guyana (Figure 5.11) (EPR = 75 – 86 m year^{-1}). Another site of consistent accretion was also detected outside of the study area in the Brazilian region of Uaça (EPR = 62 – 95 m year^{-1}) (Figure 5.11).

Shoreline locations that eroded during every time period included central and eastern stretches of Shell Beach, Guyana (Figure 5.8) (EPR = -15 – -25 m year^{-1}), Turtle Beach, Guyana (EPR = -62 – -95 m year^{-1}) and Paramaribo, Suriname (EPR = -21 - -33 m year^{-1}). The shoreline near Kourou and Cayenne, French Guyana, remained stable during every time period. Other locations which remained stable during every time period were primarily located at river inlets.

Mean EPR across Guiana did not significantly vary, and remained close to zero, between each time period (Figure 5.12). These near-zero values indicate that the coastline is in a state of equilibrium, where sediment losses and gains are balanced across the entire Guiana coastline. The smallest standard deviation in EPR value between was recorded between 1990 and 1999 and was almost 40 m less than the largest standard deviation between 2014 and 2020 (Figure 5.12 (a) – (d)).

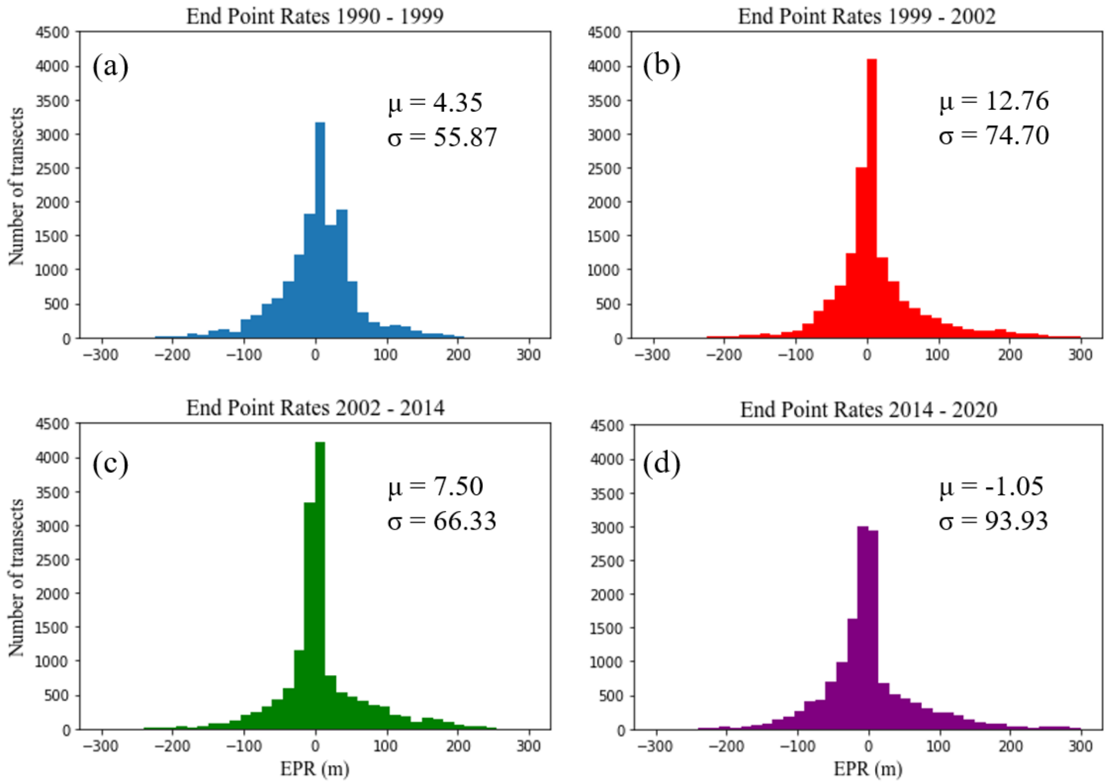


Figure 5.12: Histogram of end point rates across the Guiana coastline between: (a) 1990 and 1999, (b) 1999 and 2002, (c) 2002 and 2014, (d) 2014 and 2020. Positive and negative values correspond to eroding and accreting transects respectively.

5.3.3. Weighted Population Score

The highest Weighted Population Scores (WPS) were found in the major towns and capital cities in the region (Figure 5.13). A key component of the WPS is the number of people living within 10 km of a stretch of shoreline. Within the three countries, the maximum number of people living within 10 km of any transect was higher than 366,000 (Georgetown, Guyana), 201,000 (Cayenne, French Guyana) and 174,000 (Paramaribo, Suriname) (Figure 5.13). The number of people living within 10 km of the shoreline in South East Guyana consistently remained above 2500 people, although population densities were far lower in the north and west of the country. Fewer than 100 people lived within 10 km of 51% of the Surinamese and French Guyanese coastline, although more than 30,000 people lived within 10 km of the coastline in major towns including Nieuw Nickerie, Suriname and Kourou, French Guyana (Figure 5.13).

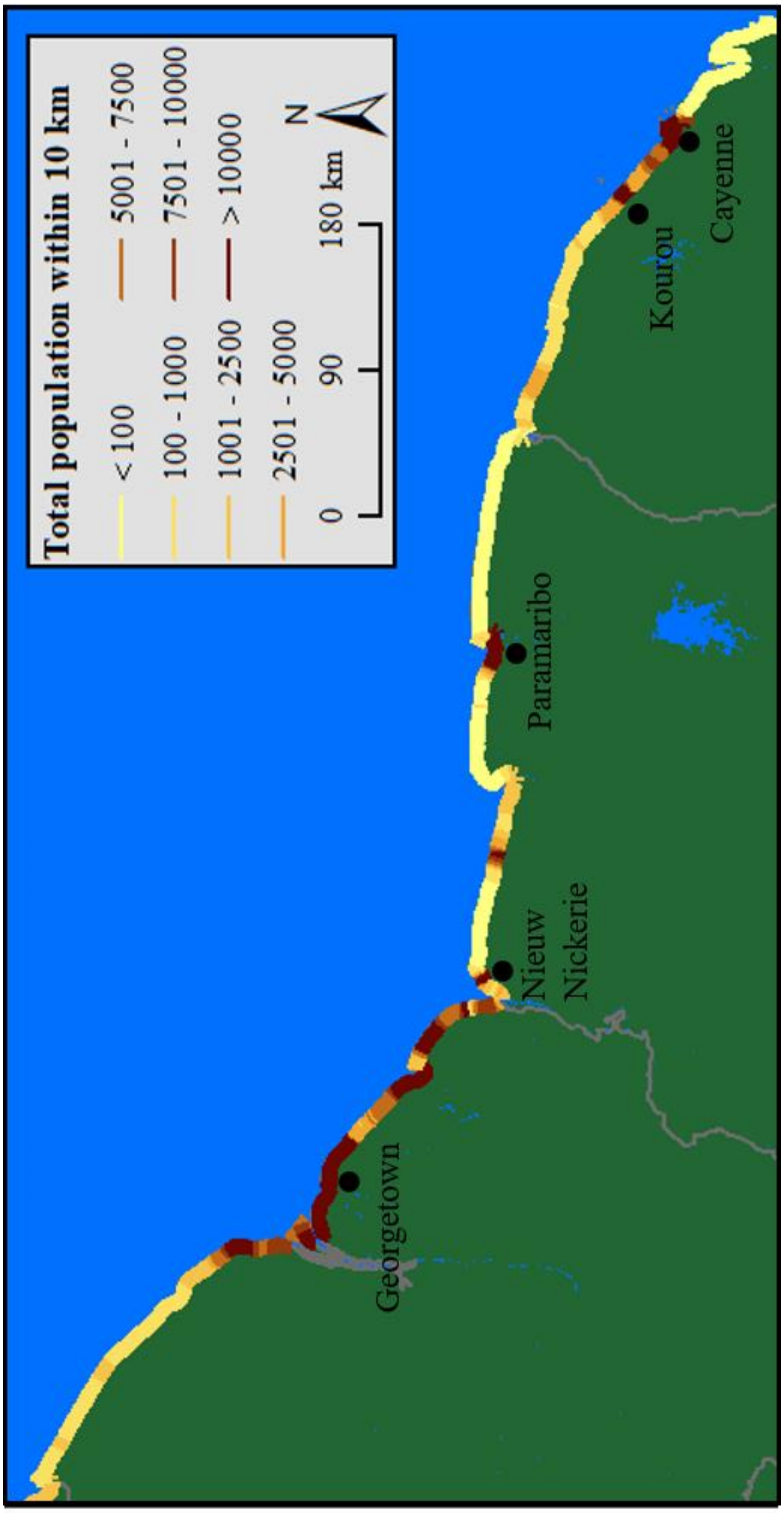


Figure 5.13: Total population living within 10 km of each transect along the Guiana shoreline. Population density transects are overlaid on a 500 m resolution Terra MODIS surface reflectance image classified into land and water classes (ORNL DAAC, 2018).

5.3.4. Coastal risk

5.3.4.1 Coastal risk scores

The two sites with the highest Risk Index were Paramaribo, Suriname (-0.92) and Mana, French Guyana (-0.81) (Figure 5.14). Despite the high population densities in cities such as Cayenne and Kourou, relatively small NSC rates contributed to a low total risk score (less than ±0.19) (Figure 5.14). Conversely, sites with relatively lower population densities but very rapid rates of erosion recorded high total risk scores including Shell Beach, Guyana (-0.48), Turtle Beach, Guyana (-0.64) and Sinnamary, French Guyana (-0.78) (Figure 5.14). The areas with the largest positive risk scores (corresponding to locations with high population density and fast rates of accretion included sites to the west of Paramaribo, Suriname (0.93) and west of Georgetown, Guyana (0.79), although the shoreline immediately adjacent to Georgetown was either stable or slowly retreating (total risk score = -0.05 – 0.36) (Figure 5.14).

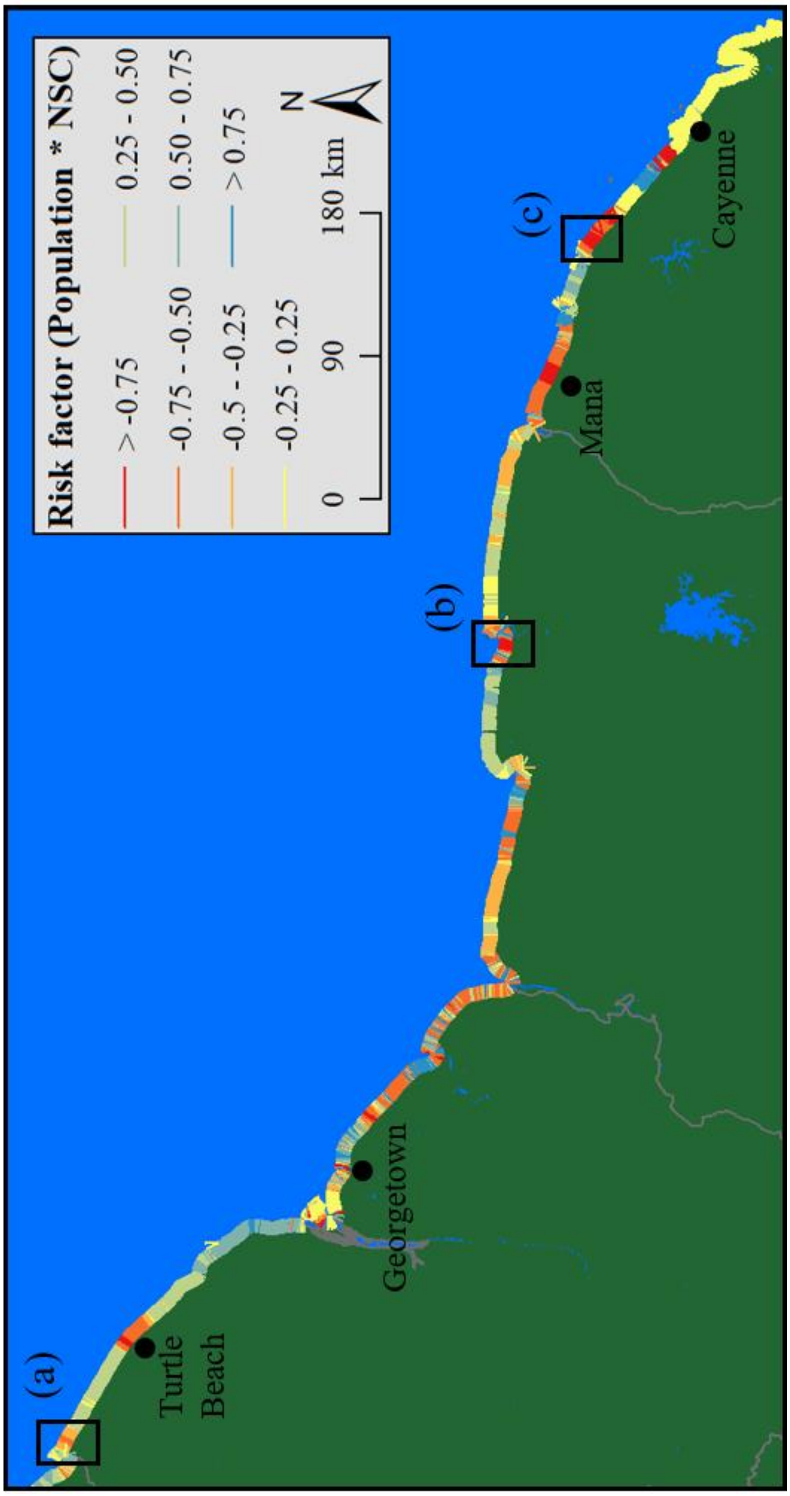


Figure 5.14: Variation in Risk Index across the Guiana coastline. Positive and negative values correspond to locations which experienced net accretion and erosion respectively between 1990 and 2020. Shoreline change rates within the three black squares at (a) Shell Beach, Guyana, (b) Paramaribo, Suriname and (c) Sinnamary, French Guyana are analysed

further below. Risk level transects are overlaid on a 500 m resolution Terra MODIS surface reflectance image classified into land and water classes (ORNL DAAC, 2018).

5.3.4.2 Annual scale shoreline dynamics

Shoreline position and rates of change were subsequently studied at three sites across the Guiana coastline with high risk index scores. These locations were: i) the site which experienced the greatest fluctuation in the rate and direction of shoreline change (Sinnamary, French Guyana); ii) the location with the highest risk index score (Paramaribo, Suriname); and iii) one location with a strong positive and negative risk index score immediately adjacent to each other (Shell Beach, Guyana). To study the dynamics of these shorelines in greater spatio-temporal detail, VEdge_Detector predicted the location of the coastal vegetation edge in these locations using annual 3 – 5 m spatial resolution Planet imagery (Section 5.2.7.1).

5.3.4.2.1. Sinnamary, French Guyana

The shoreline northeast of Sinnamary, French Guyana, was the site with the largest fluctuation in scale and direction of change during the four time periods (Figure 5.10). Using a combination of Planet and Landsat imagery, the position of the shoreline at three transect locations across the mud bank at Sinnamary was shown to oscillate between 1990 and 2021. At all three locations, the vegetation line retreated after 1990, reaching its minimum extent in 1999 (Figure 5.15). The mud bank subsequently accreted at all sites from 1999. At its westerly terminus (Figure 5.15 (a) i and (b) i), the mud bank reached its maximum extent between 2014 and 2016, with an average accretion rate of 245 m year^{-1} between 2010 and 2016. Between 2016 and 2021 the head of the mud bank retreated (mean EPR = -288 m year^{-1}). By 2021, the vegetation line was less than 400 m seawards of its position in 1990.

The oscillation in vegetation line position at the two transects at the centre and eastern edge of the mud bank were four years in front of the oscillation at the westerly mud bank head (Figure 5.15 (a) ii, (a) iii, (b) ii and (b) iii). At sites (ii) and (iii), the vegetation line reached its maximum extent in 2010, before retreating every year between 2010 and 2021 with mean erosion rates of -214 m year^{-1} and -255 m year^{-1} respectively (Figure 5.15 (a) ii, (a) iii, (b) ii and (b) iii). The shorelines at site (ii) and (iii) retreated at slower rates of -56 m $year^{-1}$ and -1 m $year^{-1}$ respectively between 2019 and 2021. By 2021, the vegetation line at sites (ii) and (iii) had retreated to a position landwards of the previous minimum extent recorded in 1999.

Changes in EPR values at all three transects took the approximate form of a sine curve. The smallest difference in EPR values between years occurred when the mud bank was close to its maximum and minimum extent. Changes in EPR were much larger when the mud banks were at an intermediate extent (Figure 5.15).

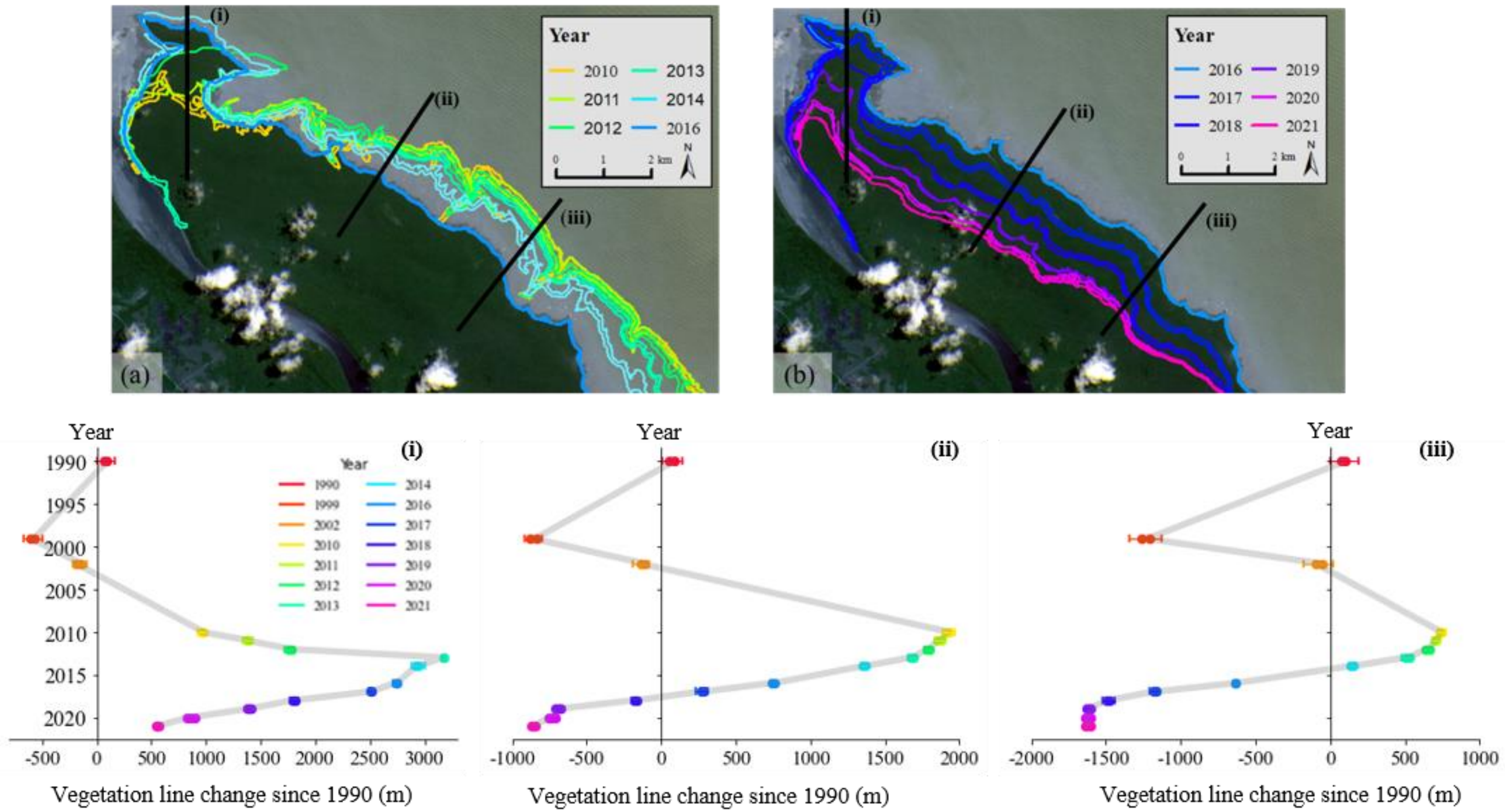


Figure 5.15: Vegetation line position identified by VEdge_Detector at Sinnamary, French Guyana between (a) 2010 and 2016, (b) 2016 and 2021. Shorelines are represented by the confidence contours derived by VEdge_Detector. Shoreline positions are overlaid on the 2016 5 m RapidEye image in both figures. Changes in shoreline position along three perpendicular transects (i), (ii) and (iii) are shown in the inset graphs. For all three inset graphs, 0.0 is the location of the seaward 0.1 confidence contour in 1990. The solid circles and vertical bars represent the location of the landward and seaward 0.9 and 0.1 confidence contours respectively. The shoreline change maps and graph insets use the same colour ramp.

5.3.4.2.2. Paramaribo, Suriname

Between 1990 and 2020, the vegetation extent at the western edge of the Paramaribo study site displayed a similar oscillating pattern, as shown for Sinnamary (Figure 5.16 (i)). The shoreline accreted between 1990 and 2002, reaching its maximum extent in 2002 (mean EPR = 67 m $year^{-1}$), before eroding between 2002 and 2020 (mean EPR = -62 m $year^{-1}$). The rate of retreat slowed between 2016 and 2020 (EPR = -18 m $year^{-1}$). By 2020, the coastal vegetation edge position was less than 150 m from its position in 1990.

In contrast, the shoreline immediately adjacent to the city limits of Paramaribo retreated during every time period between 1990 and 2020 (mean EPR = -38 m $year^{-1}$) (Figure 5.16 (iii)). The erosion has resulted in the formation of new headlands containing high density buildings (Figure 5.16 (a)). In 1990, the distance between the road shown in Figure 5.16 (a) and the coastal vegetation edge was more than 900 m. By 2020, more than 2.5 km of this road was situated less than 110 m from the coastal vegetation edge.

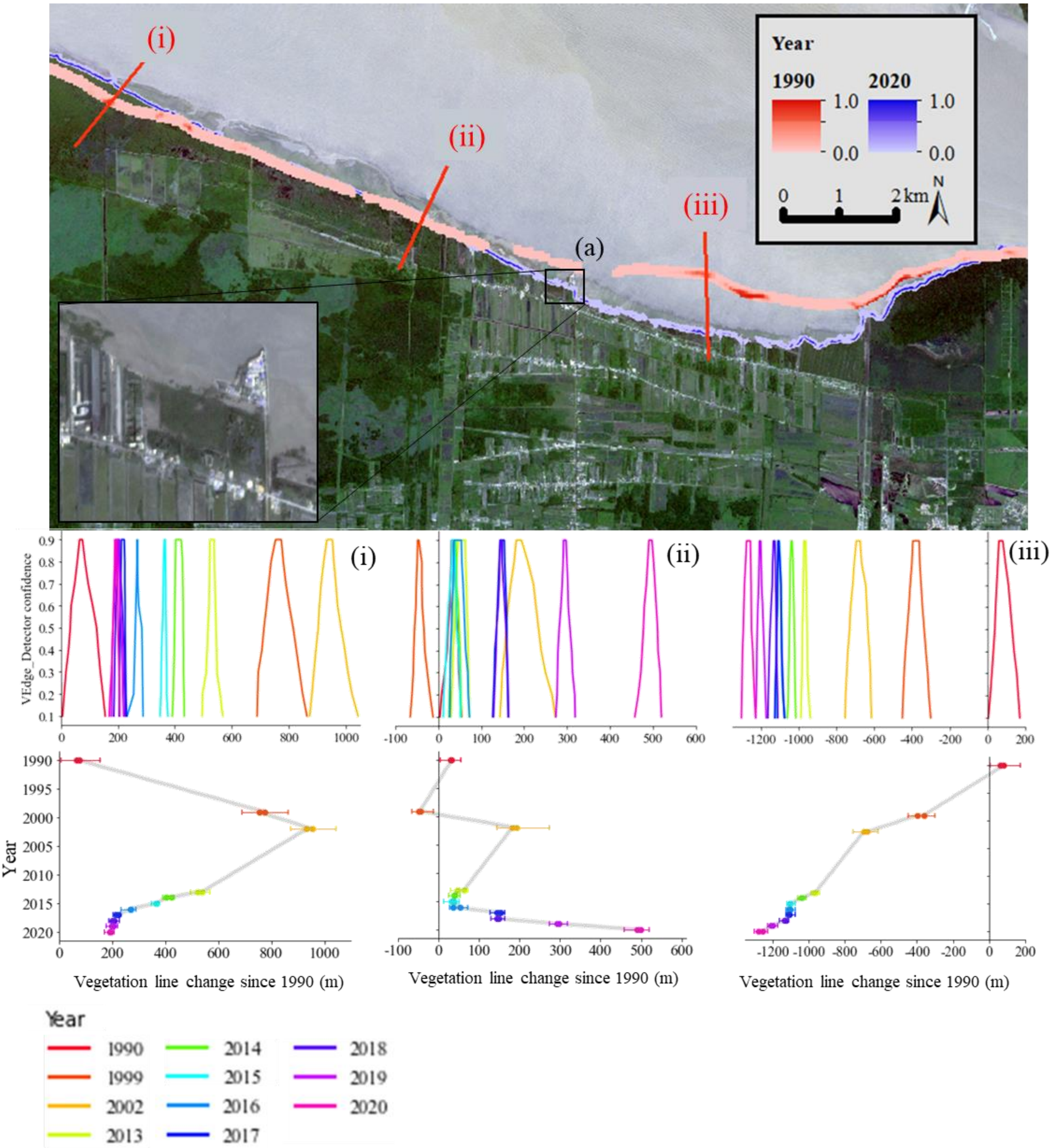


Figure 5.16: Vegetation line position at Paramaribo, Suriname, identified by VEdge_Detector in 1990 (red) and 2020 (blue), overlaid on the 3 m spatial resolution Planetscope image from 2020. Change in shoreline position along three perpendicular transects (i), (ii) and (iii) are shown in the inset graphs for 1990, 1999, 2002 and every year between 2013 and 2020. For all three inset graphs, 0.0 is the location of the seaward 0.1

confidence contour in 1990. The solid circles and vertical bars represent the location of the landward and seaward 0.9 and 0.1 confidence contours respectively. Inset (a) shows the position of a headland containing a high density of buildings and coastal road at a larger scale

5.3.4.2.3. Shell Beach, Guyana

The shoreline at the western terminus of the Shell Beach mud bank accreted every year between 2010 and 2020 (mean EPR = 104 m year^{-1}). In contrast, the vegetation line at the eastern end of the mud bank eroded every year between 2010 and 2020 (mean EPR = 101 m year^{-1}) (Figure 5.17). Between 2011 and 2016, the shoreline immediately adjacent to Shell Beach village, in the centre of the mud bank, remained stable (mean EPR = -2.5 m year^{-1}) and in some years the shoreline accreted (2016 – 2017 EPR = 16 m year^{-1}) (Figure 5.17 (a) – (d)). From 2017 to 2020, the shoreline at Shell Beach eroded and the entire village was lost to the sea (mean EPR = -71 m year^{-1}) (Figure 5.17 (e) and (f)). The most westerly transect that recorded a negative EPR (net erosion) value greater than -20 m year^{-1} consistently migrated towards and beyond Shell Beach village between 2010 and 2020 (Figure 5.17 (a)). The distance between Shell Beach and the nearest transect to the east of Shell Beach with an EPR greater than -20 m year^{-1} reduced from 5980 m in 2011, to 269 m in 2016. Between 2017 and 2020, this transect migrated from 397 – 1611 m to the west of Shell Beach (Figure 5.17).

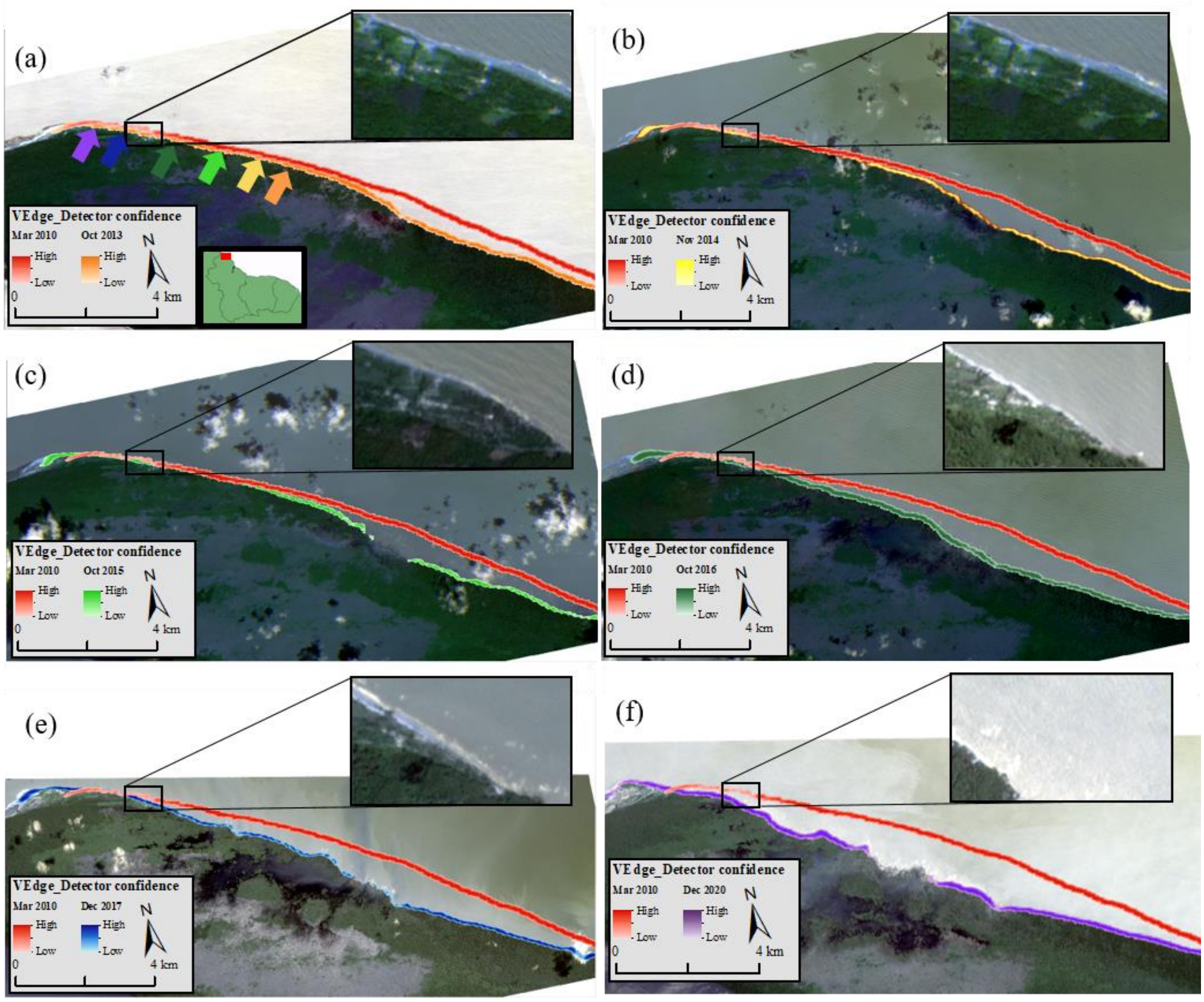


Figure 5.17: Vegetation line position identified by VEdge_Detector at Shell Beach, Guyana, in 2010 (red) compared with (a) 2013; (b) 2014; (c) 2015; (d) 2016; (e) 2017; (f) 2020. The insets show the shoreline position at Shell Beach village at a larger scale. The arrows in (a) show the closest transect to the east of Shell Beach which recorded an EPR greater than -20 m year^{-1} for each of the years. The colours of the arrows correspond to the colour ramp for the corresponding year.

5.3.5. Statistical relationships between EPR and the NAO and ENSO indices

During years when the DJFM NAO index was strongly positive, the mud bank at Sinnamary retreated. Conversely, when the DJFM NAO index was weakly positive or negative, the mud bank at Sinnamary accreted. The maximum rate of erosion (EPR = -267 m year^{-1}) was recorded in 2018, after four consecutive years with DJFM NAO index greater than 1.25 (Figure 5.18). The Sinnamary mud bank retreated in most years with a negative ENSO index (El-Niño), and accreted in years with a positive ENSO index (La-Niña).

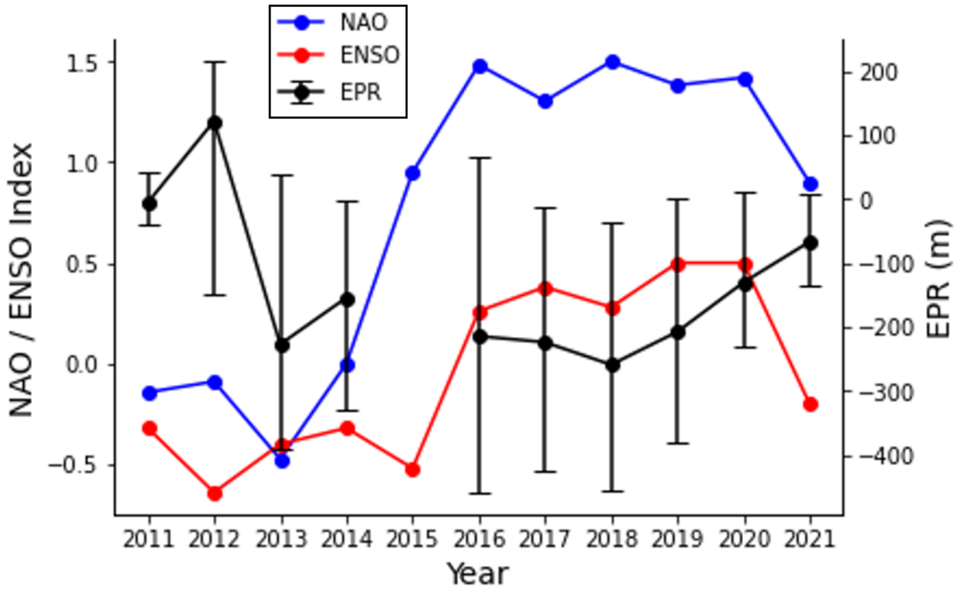


Figure 5.18: Comparison or NAO index, ENSO index and rates of shoreline change. Five-year average December January February March North Atlantic Oscillation (DJFM NAO) index (blue), ENSO index (red) and average end point rates (EPR) (black) along the Sinnamary coastline between 2011 and 2021. Black dots show mean EPR rates, the lower and upper bar extents show the 10 and 90 percentile values respectively. EPR data is missing for 2015 when no Planet image was available.

Transect (ii) at Paramaribo was the only site containing non-normally distributed EPR values. At this site, the Spearman's Rank correlation coefficient was calculated whilst at all other sites the Pearson's correlation coefficient was calculated. A statistically significant correlation between EPR and both the NAO and ENSO indices existed at transect (i) and (ii) at Sinnamary. At transect (iii) at Sinnamary, EPR was only statistically significantly correlated with NAO (Table 5.1). At all transects at both sites, a negative correlation existed between EPR and both the NAO and ENSO indices. The strongest negative correlation existed between EPR and NAO at transects (i) and (ii) at Sinnamary (r = -0.84 and -0.64 respectively). At Paramaribo, no statistically significant correlation between EPR and either the NAO index or the ENSO index existed at any transect (Figure 5.19 (d) – (f); Table 1). Correlation coefficient values were consistently weak, less than ±0.41, at all transects at Paramaribo. No statistically significant correlation between EPR and MHWL existed at any site.

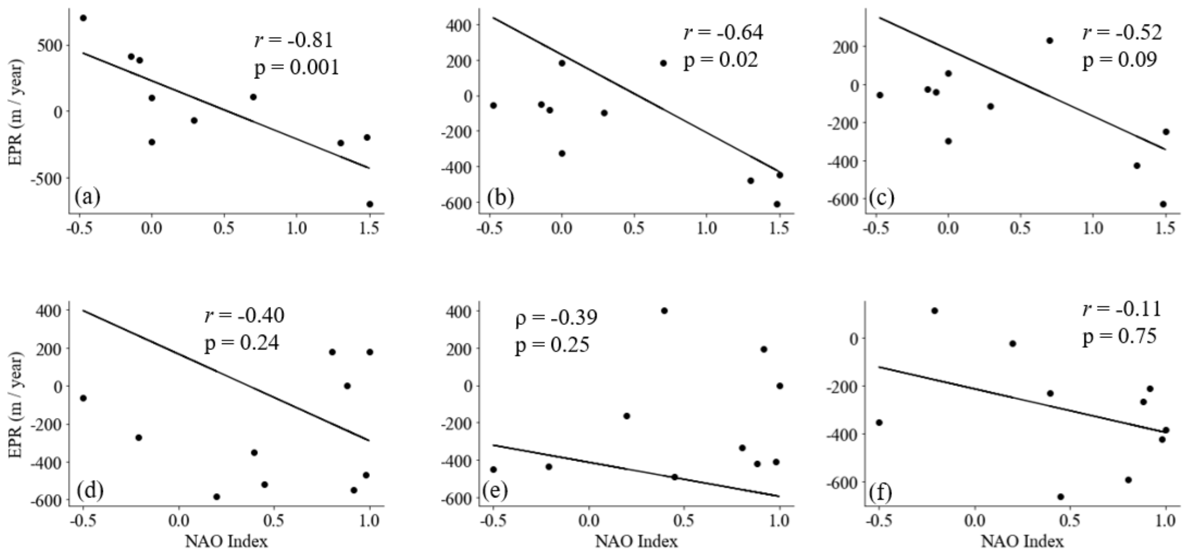


Figure 5.19: Scatter plots of EPR verses NAO index for (a) – (c) transects (i), (ii) and (iii) at Sinnamary (see Figure 5.14) and for (d) – (e) transect (i), (ii) and (iii) at Paramaribo (see Figure 5.16). For (a) – (d) and (f) the Pearson's correlation coefficient, *r*, and associated p-value is provided. For (e), the Spearman Rank's coefficient, ρ, and associated p-value is provided. Note: different scales are used on the y-axis for each transect.

Table 5.1: Pearson's correlation coefficients, r, between EPR and (a) NAO index, (b) ENSO index and (c) MHWL at Sinnamary and Paramaribo. At transect (ii) at Paramaribo, the Spearman's Rank correlation coefficient was calculated because the EPR data at that transect was not normally distributed.

	Sinnamary						Paramaribo					
	Transect (i)		Transect (ii)		Transect (iii)		Transect (i)		Transect (ii)		Transect (iii)	
	r	p	*r*	p	*r*	p	*r*	p	*ρ*	p	*r*	p
NAO	-0.84	0.001	-0.64	0.02	-0.52	0.04	-0.4	0.24	-0.39	0.25	0.11	0.75
ENSO	-0.81	0.001	-0.61	0.03	-0.45	0.14	0.01	0.96	0.05	0.88	0.41	0.24
MHWL	-0.06	0.85	0.68	0.12	0.52	0.11	0.41	0.31	0.52	0.18	-0.37	0.36

5.3.6. Regional scale dynamics

5.3.6.1. Correlations with Sinnamary site

Across the Guiana coastline, there were four major clusters of transects where EPR rates were strongly positively correlated with the EPR values recorded at Sinnamary (Figure 5.20). These stretches of shoreline were located near Mana, French Guyana, the Wia-Wia and Coppername nature reserves in Suriname and stretches of coastline between Nieuw Nickerie and Totness, Suriname. No cluster of transects, determined here as ≥ 5 adjacent transects, were correlated to the Sinnamary site along the whole of the coastline of Guyana. The four main clusters of transects which correlated with Sinnamary were consistently spaced between 180 and 190 km apart.

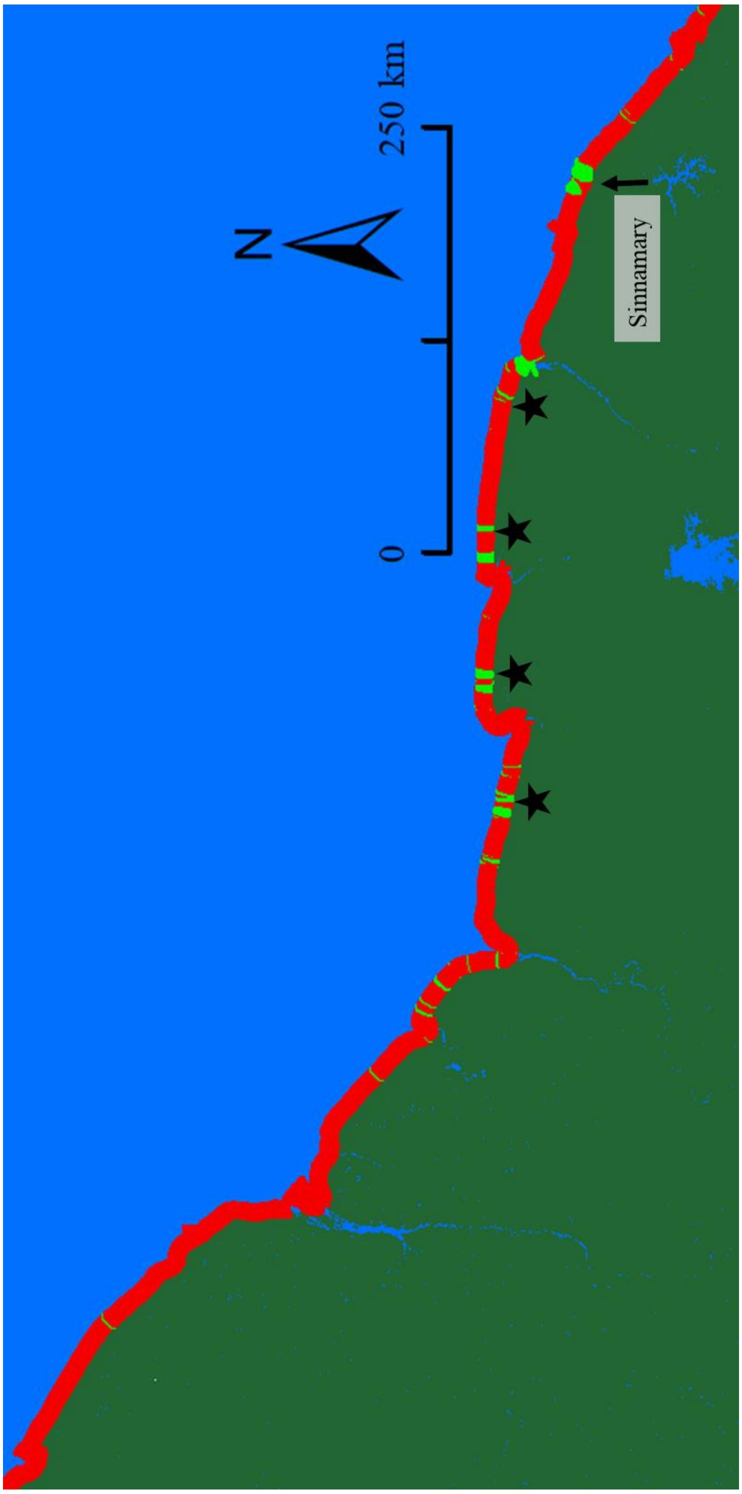


Figure 5.20: Locations (in green) where end point rate (EPR) values had a strong positive correlation with EPR values at Sinnamary, French Guyana. Transects are coloured green when correlation coefficient values, r, are greater than 0.9. Black stars correspond to stretches of shoreline consistently positioned 180 – 190 km apart, where transect EPR values are correlated with the Sinnamary site. The Sinnamary site itself is denoted by the black arrow.

5.3.6.2. Mana

By producing a line graph of EPR values across a 70 km stretch of shoreline near Mana, during each of the time periods, it was possible to deduce a westerly migrating erosional hotspot (Figure 5.21, positive values correspond to shoreline erosion). Between 1990 and 1999, the highest EPR values were found to the east of the study site, adjacent to Organabo. The location with the peak rates of erosion subsequently migrated westwards over the subsequent three time periods. Between 2014 and 2020, the peak in erosion rates was located less than 5 km to the east of Mana (Figure 5.21). The pattern for accretion (negative values in Figure 5.20) was less distinct. However, at Organabo, after the high rates of erosion between 1990 and 1999, the shoreline consistently accreted during each subsequent time period.

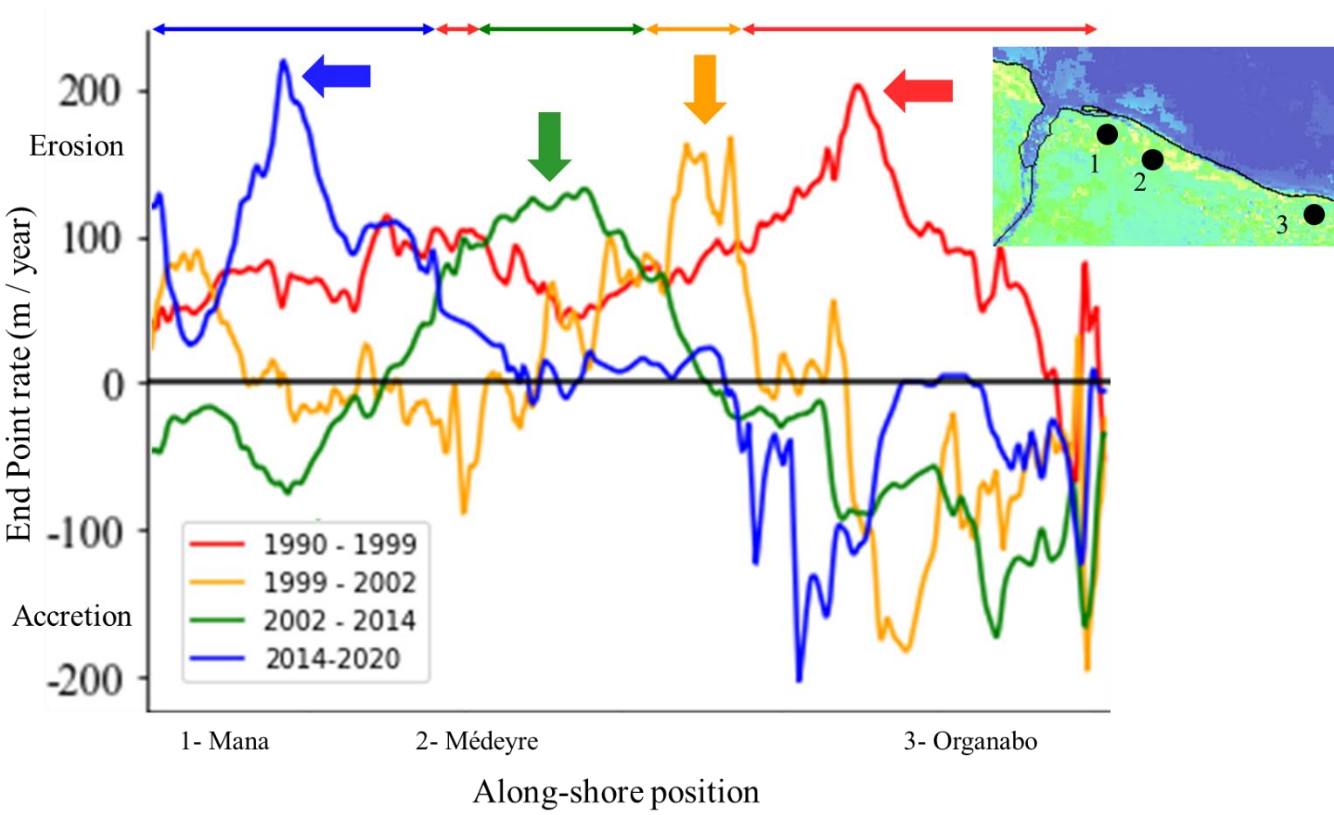


Figure 5.21: Comparison of EPR values for each of the four time periods within the study near Mana. The x-axis corresponds to alongshore position between Mana and Organabo, corresponding to approximately 700 transects across the study site. The map represents the stretch of coastline studies, with the relative position of Mana, Médeyre and Organabo

numbered 1, 2 and 3 respectively. The vertical or left pointing arrows emphasise the peak locations with maximum rates of erosion during each of the four time periods. The horizontal bars at the top of the graph represent the time period during which rates of erosion were greatest for those transects. Positive and negative values correspond to erosion and accretion respectively.

5.4. Discussion

Applications of ML tools to detect shoreline position and associated rates of change have typically been restricted to relatively short, less than 100 km, stretches of coastline (Goldstein et al., 2019; Elnabwy et al., 2020; Rogers et al., 2021). Despite being insightful for understanding local scale dynamics, this Chapter has demonstrated the potential of applying ML techniques to detect shoreline position at a supra-national scale, with an ability to identify patterns of coastal morphodynamics that are not detectable in local scale studies. These morphodynamics included the identification of spatially distinct stretches of shoreline that respond in similar ways to the same extraneous forcing factors (Figure 5.20), migrating erosional hotspots (Figure 5.21), and locations where the greatest number of people are exposed to shoreline change (Figure 5.14). This regional-scale information may assist both national governments and international NGOs to identify locations most exposed to current and future shoreline change.

Due to the inability to access this study site, it was not possible to collect ground-referenced measurements to ascertain error terms generated in this study. Chapter 4 determined that VEdge_Detector identified the shoreline position with a root mean square error of less than 6 m (two image pixels) at the majority of sites. To account for the coarser Landsat imagery used in this study, it is estimated that mean error terms in this study probably range between 30 – 60 m (corresponding to between one and two image pixels). Further investigation could substantiate these error terms via the application of ground- referenced studies.

5.4.1 Shoreline response to extraneous forcing factors

Mean EPR values across the Guiana shoreline were near-zero between each of the four time periods when Landsat imagery was available (Figure 5.12, EPR 1990 – 1999 μ= 4.35 m year^{-1}, EPR 2014 - 2020 μ= -1.05 m year^{-1}), showing that there was no net trend in erosion or accretion during this period. The increasing trend in standard deviation values over time (EPR 1990 – 1999 σ = 55.87 m year^{-1}, EPR 2014 - 2020 σ = 93.93 m year^{-1}) reveals that a greater number of transects were advancing or retreating by more than 60 m year^{-1} by 2020. This indicates no net sediment loss or gain across the entire Guiana coastline, but local-scale (10^{-1} – 10^{1} km) stretches of shoreline are becoming more dynamic, increasing the total amount of erosion during interbank phases (Figure 5.15). Greater rates and magnitude of shoreline change may impact upon local ecological and human receptors. Greater EPR rates along mud banks in French Guiana have, for example, a deleterious effect on sea turtle populations by flooding their nesting sites (Caut et al., 2010). Further, human settlements that were previously landwards of the minimum extent of mud banks may become more exposed to coastal erosion during future mud bank oscillations (Bhargava et al., 2021). Consideration of this increase in the amplitude of mud bank oscillations may also be beneficial when planning and positioning future human coastal developments.

Sinnamary, French Guyana, was the most dynamic stretch of shoreline across the entire Guiana coastline, where it was identified that vegetation edge position oscillates over time (Figure 5.9). A strong negative correlation between EPR values at Sinnamary and the NAO (r = -0.84, p= 0.001) and ENSO index (r = -0.81, p= 0.001), indicates that NAO and ENSO may have a strong influence on the mud bank and mangrove extent (Figure 5.15, Figure 5.18, Figure 5.19, Table 5.1). Between 2013 and 2020, rapid retreat (mean EPR greater than -100m year^{-1}) coincided with years with consistently positive NAO index values (Figure 5.18), with three of the ten highest positive NAO values recorded between 2013 and 2020 (NOAA, 2021). Reducing EPR values between 2018 and 2020 also coincided with weaker NAO and ENSO indexes (Figure 5.18). Other studies have found that stronger trade winds (present during years with a strong positive NAO index) can generate larger significant wave heights, promoting greater erosional change along the Guiana coastline (Young et al., 2011; Young et al., 2019). ML and numerical modelling methods are becoming increasingly proficient at predicting NAO index values in future years (Wang et al., 2017; Yuan et al., 2019). Increased understanding of the relationship between EPR rates and NAO index, combined with increased NAO index predictive capability, could provide a greater predictive capacity to forsee years when the greatest rates of coastal change may occur.

No statistically significant relationship between Mean High-Water Level (MHWL) and EPR was identified at any Sinnamary transect (Table 5.1). Whilst peak MHWL in 1999 coincided with minimum mangrove extent at Sinnamary, the next peak in MHWL occurred in 2010, which coincided with the maximum mangrove extent (Figure 5.15 (a) – (c)). In contrast, previous studies have primarily attributed the change in mangrove extent at Sinnamary to fluctuations in the Mean High-Water Level (MHWL), driven by the 18.6-year lunar nodal cycle (Gratiot et al., 2008; Anthony et al., 2014). These previous studies used the waterline as a proxy of shoreline position, and the discrepancies between the findings presented here and those in previous studies may be due to the close relationship between waterline position and MHWL, particularly along shallow-gradient mud bank coastlines, where it is difficult to tidally correct waterline position (Bhargava et al., 2021). Despite knowledge of its limitations (Burningham and French, 2013), the NAO index may prove to be a better predictor of rates of change to the coastal vegetation edge position and corresponding risks of coastal flooding and erosion to adjacent receptors. Future work should continue to collect detailed data pertaining wind and wave conditions, when monitoring mud bank extent at Sinnamary, and identify if the two shoreline proxies continue to exhibit different relationships to the extraneous forcing factors.

Correlation analysis enabled the identification of four spatially distinct clusters of transect locations where the shoreline position changed in a similar magnitude and direction to Sinnamary between each of the time periods (Figure 5.20). These stretches of coastline are consistently spaced approximately 180 – 190 km apart, which is consistent with previous assertions that there is a consistent spacing between the dynamic mud banks across the Guianas (Allison et al., 2000; Anthony et al., 2010). Large proportions of the Guiana Coastline are remote and inaccessible (Gardel and Gratiot, 2005), and the idenfitication of spatially distinct stretches of shoreline that have similar rates of shoreline change could aid the targeting and optimisation of future monitoring projects to fewer locations. However, as discussed below, the coastline of Guyana did not contain any stretches of shoreline that exhibited similar dynamics to those at Sinnamary.

The coastlines that exhibit dynamics like those at Sinnamary are all considered to be relatively undisturbed stretches of shoreline, where few human shoreline interventions have been implemented (Figure 5.20). No transects in Guyana had EPR values that correlated with those at Sinnamary, where a sea wall has been constructed along nearly the entirety of the country's coastline (Vaughn, 2017). Likewise, at Paramaribo, Suriname the site with the highest risk index across the entire Guiana coastline, no statistically significant correlation existed between EPR and either the NAO or ENSO index (Figure 5.19; Table 5.1), despite the presence of a mud bank with similar size and orientation to Sinnamary. The weakest correlation between EPR and NAO was identified at transect (iii), where the shoreline consistently retreated between each time period (Figure 5.16). EPR decreased between 2015 and 2017 (EPR = 8 m year^{-1}), suggesting that the site was reaching a minimum value in shoreline oscillation (Figure 5.16 (iii)). EPR began to accelerate again, however, between 2017 and 2020 (EPR = 50 m year^{-1}), despite the NAO index decreasing by 2020 (Figure 5.18); indicating that other processes, suggested here to be human shoreline modifications, were controlling EPR values at this site.

The decoupling of EPR rates to the NAO and ENSO index and cessation of the 30-year cycle in shoreline position at transect (iii), Paramaribo and across Guyana, may be related to the extensive clearance of mangrove forest for agricultural production and construction of sea walls along these stretches of shoreline (World Bank, 2018, de Jong et al., 2021). At Paramaribo, along more than 6 km of the shoreline adjacent to transect (iii), mangrove forest has been removed and agricultural land is situated immediately adjacent to the shoreline (Figure 5.16 inset (a)). Hard engineered sea defences have also been constructed around buildings, forming abrupt headlands with high housing density valued at an estimated US$ 86 million (Figure 5.16 inset (a); Burke and Ding, 2016). Sea walls have in some places enclosed mangrove forests, preventing propagules from being transported from these forests to establish on, and stabilise, new mud banks (Anthony and Gratiot, 2012). Without the establishment of new mangrove vegetation, mud banks are more exposed to wave action, leading to greater mud bank liquefaction and disintegration (Lewis, 2005; Anthony and Gratiot, 2012).The removal of mangrove forest is another route, alongside sea wall construction, by which propagules cannot establish on newly formed mud banks, to establish their stilted rooted system, trap sediment and reduce incident wave power (Anthony and Gratiot, 2012; Winterwerp et al., 2007 & 2013; Brunier et al., 2019). To provide further evidence for this relationship between mangrove clearance and continued shoreline retreat, more detailed assessment, combining local knowledge, is necessary to understand whether mangrove clearance preceded the continuous erosion now present. This information could support the implementation of management options, such as preventing further mangrove clearance (de Jong et al., 2021), to improve the likelihood of the mud banks' continued alongshore migration.

5.4.2. Shifting erosional hotspots

By identifying the seaward vegetation extent at a supra-national scale between multiple times dates, it was possible to determine shifting erosion hotspot features and thus potentially predict locations likely to experience increased rates of shoreline retreat in the future. The largest scale shifting erosion hotspot was identified near Mana, French Guyana where the transects recording the maximum rates of shoreline retreat moved westwards by nearly 55 km between 1990 and 2020 (Figure 5.20). As of 2020, this erosional hotspot was located approximately 5 km to the east of Mana, meaning that at current rates, the erosional hotspot will be immediately adjacent to the town by 2025. This information could be useful for town planners, home and business owners and other stakeholders in the region, to ensure that further development does not take place 'in the path' of this erosional hotspot.

A shifting erosional hotspot was also located at Shell Beach, Guyana (Figure 5.17). The transect nearest to Shell Beach village with an erosional EPR value greater than 20 m $year^{-1}$ reduced from 5980 m in 2011 to 269 m in 2016 (Figure 5.17 (a)). This consistent westerly migration in the site of erosion could have provided a warning to local governments and to the residents and business owners of Shell Beach of the likelihood that the position of Shell Beach village would be subject to coastal erosion. In contrast, examination of the Planet satellite imagery shows that the size of the Shell Beach settlement continued to expand between 2010 and 2016 (see insets Figure 5.17 (a) – (d)). This shifting erosional hotspot led to the forced evacuation of Shell Beach village in 2017, consisting of more than 50 households and numerous tourism businesses (Kaieteur News, 2017; Staebrook News, 2017). Future research could use automated methods, such as VEdge_Detector, on higher temporal resolution Planet imagery to qualitatively identify other stretches of shoreline in the 'path' of the migrating erosion hotspot, such as Mahaicony, Guyana (Figure 5.9), which could then be validated using ground-referenced surveys and local knowledge. This could aid better planning in these locations, including restricting further development in particular areas, or providing enough advanced warning to enact a planned evacuation from these stretches of shoreline.

5.4.3. Identifying populations exposed to shoreline change

Locations where the highest number of people are exposed to shoreline change were identified by combining shoreline change data with population indices, obtained from WorldPop (Stevens et al., 2015; Figure 5.14). Sites with high absolute exposure values included Paramaribo, Suriname and Mana, French Guyana, and sites were also identified with high negative risk index values immediately adjacent to sites with high positive values, e.g. Shell Beach, Guyana and Mahaicony, Guyana (Figure 5.14). Previous studies have assessed the Guiana coastline to be subject to relatively low risks to coastal hazards (Calil et al., 2017), although shoreline change metrics were not included in total hazard scores, limiting the ability of the study to claim it is generating total risk scores. This highlights the importance of considering rates of shoreline change when comparing relative levels of coastal risk across an area of interest.

To develop from exposure to risk scores, population dynamic and shoreline change values require integration with data pertaining to other vulnerability and exposure indices (Kron, 2013). Datasets relating to land cover, the location of ecologically protected areas, as well as nationally and internationally significant infrastructural sites located outside of major settlements, including roads, airports and power stations, are necessary to make a fully informed assessment about the relative levels of exposure receptors face to current and future shoreline change. Receptor vulnerability could be considered using datasets on housing quality, levels of deprivation and the location of informal settlements. Despite their importance in moderating shoreline dynamics along the Guiana coastline, a publicly available dataset pertaining to the location and type of human shoreline modification factors employed across the Guiana coastline could not be found. Collaboration between local communities, research groups and stakeholders with knowledge on shoreline intervention position could aid the development of such as dataset, which could better inform future studies on how human interventions are perturbing dynamics across the entire Guiana coastline.

Chapter 6. Potential developments of machine learning to shoreline change and coastal risk

This chapter explores multiple concepts and discoveries regarding ML and its application to coastal risk, focussing on the ability to use ML to further our understanding of shoreline position and coastal change. A comparison of the different ML and non-ML tools applied in this thesis is explored further. These tools were trained and developed to exclusively detect the coastal vegetated edge via multispectral remote sensing imagery. A deeper understanding of the relative performance of the different techniques is necessary to inform future decisions on the most appropriate vegetation edge detection tool to use by other researchers and stakeholders in the coastal zone, and whether these tools require any further development or refinement. This decision may also vary depending upon the shoreline type, position and scale present in the imagery, as well as the coastal zone dynamics of interest.

Tool performance, in terms of accuracy, specificity and the ability to generalise to different coastlines is a key component of tool selection. The future selection of shoreline detection tools will commonly consider other aspects, most notably with the rise of Big Data and the frequent capture of global scale remote sensing imagery, the processing speed of the tool, and the estimated time required to detect the shoreline position will become increasingly important. When multiple tools have similar performance and computational speed, tool selection may be determined by their ease of development, interpretability, and their transferability to a range of remote sensing imagery with varying spatial resolution.

Shoreline proxies, which can be extracted from multispectral remote sensing imagery, each provide a different representation of coastal zone processes (Chapter 4). Irrespective of the performance and attributes of vegetation edge detection tools, they will only be valuable to coastal stakeholders if the coastal vegetation edge is a suitable shoreline proxy for discerning the coastal dynamics of interest. Coastal stakeholders will commonly be interested in utilising the shoreline proxy which best represents how exposed a coastal receptor is to shoreline change. The multifaceted nature of coastal risk means shoreline detection tools cannot be used in isolation to determine relative coastal receptor exposure to shoreline change. To derive exposure and risk values, shoreline detection tools require integration with other tools pertaining to different aspects of coastal risk. The potential to integrate multiple ML tools to ascertain relative levels of risk in the coastal zone needs to be explored in the context of the spatio-temporal resolution and spatial scale at which relative exposure values can be derived.

6.1. Is machine learning the way forward?

When selecting a tool to extract a shoreline proxy from multispectral remote sensing imagery, some of the key stakeholder questions, alongside tool performance, may include: i) 'how fast will the tool/s detect the shoreline position?'; ii) 'what training, development and computer programming requirements will be needed to set up the tool/s?'; iii) 'how easy will it be to explain, in non-technical terms, how the tool derived its outputs?'; and iv) 'will the tool work on multiple types of remote sensing imagery with different spatial resolutions?'. By considering these questions, and tool attributes, a more informed decision-making process will be engaged as to whether or not machine learning (ML) tools, such as Convolutional Neural Networks (CNN) or Support Vector Machines (SVM), will be able to provide new insights into coastal risk dynamics. And if new insights do appear likely, how will they compare with using more established, non-ML, techniques?

6.1.1 How does tool performance compare?

The unprecedented availability of global-scale remote sensing imagery makes the use of manual methods less viable, increasing the need for automated tools to extract shoreline position (Chapter 1; Figure 1.2; Gorelick et al., 2017; Tamiminia et al., 2020). Three aspects of tool performance are important during tool selection: i) tool accuracy, in terms of precisely identifying the position of the shoreline feature; ii) tool generalisability, or whether the tool can robustly detect the coastal vegetation edge in a range of coastal locations, and iii) tool specificity, or the ability to distinguish between the coastal vegetation edge and other boundaries present within the image.

6.1.1.1. Tool accuracy and generalisability

Tool accuracy is an essential consideration when justifying the use of automated methods over more traditional ground-referenced surveys or manual techniques. If error and uncertainty values are similar or larger in value than the amplitude of the shoreline dynamics of interest, it may be difficult to separate the shoreline change signal from noise, limiting the use of a tool's output (Thieler and Danforth, 1994; Pardo-Pascual et al., 2012). The ability of the tool to generalise to detect the shoreline position in a particular location will also be a key consideration, particularly if the tool has not been applied in that location previously. The accuracy of CNN and SVM and seven non-ML tools in identifying the position of the coastal vegetation edge was separately determined throughout this thesis (Chapter 3; Chapter 4), but further examination is presented here on their relative accuracy and generalisability to inform future tool selection.

The accuracy of all the tools applied in this thesis was compared across a range of coastal environments, using images of a sandy-shingle beach and dune system (Dunwich), a gravel-barrier island with salt marsh vegetation (Blakeney Point) and two separate stretches of soft-rock cliff, with cliff-top agricultural land cover (Covehithe and Holderness). VEdge_Detector, a CNN, produced the smallest RMSE (< 8 m) at every site, although NDVI threshold contouring, CoasTool and SVM identified the shoreline with similar RMSE (< 11.5 m) at Dunwich and Covehithe (Table 3.4; Table 4.2; Table 4.3). The high performance of all of these tools at Dunwich and Covehithe was attributed to the relatively short and straight coastlines in these locations, along with relatively uniform vegetation cover characteristics (Section 3.4). Images of these types of locations are commonly characterised by a consistent alongshorebut abrupt cross-shore change to the pixel values at the coastal vegetation edge. This characteristic has previously been identified as a key determiner of the performance of threshold contours (Zhao et al., 2008; Bishop-Taylor et al., 2019) and SVM (Choung and Jo, 2017) in instantaneous waterline detection. Greater differences in the performance of the tools was identified at sites containing more complex coastal vegetation boundaries.

The performance of all non-ML tools and SVM was inconsistent along shorelines with graded or discontinuous vegetation edges and shorelines with heterogeneous vegetation properties (Table 3.4). CoasTool, SVM and NDVI threshold contouring produced RMSE values greater than 34 m at Hornsey and Blakeney Point. High standard deviations (σ) in RMSE values demonstrated how these tools accurately identified the shoreline position in some locations but along other frontages the tools' outputs deviated far from the manually digitised shoreline (Figure 3.5; Figure 3.13). This inconsistency in tool performance can be attributed to the graded salt marsh vegetation found at these sites, including *Salicornia* spp., which forms isolated clumps at its seaward limit (Möller et al., 2006). This patterning results in individual remote sensing pixels containing vegetation, exposed substrate and water (Klemas, 2013; Medina Machín et al., 2019). This highlights the inability of currently available non-ML and SVM to overcome intra-site variability in the spectral properties of vegetation, caused by variability in plant species, composition and phenology, alongside environmental factors including soil moisture content and mineral composition (Belluco et al., 2006; Rahman et al., 2011; Gandhi et al., 2015; Liu et al., 2020).

In comparison, CNNs provide promise in being able to generalise to larger scale stretches of shoreline consisting of heterogeneous vegetation species. Nevetheless fragmented vegetation lines remain problematic. A continuous coastal vegetation edge was identified when traversing across vegetation boundaries, including between salt marsh and agricultural land at Blakeney Point (Figure 4.5) and mangrove forest and agricultural land at Paramaribo, Suriname (Figure 5.13). The ability to traverse different vegetation species is an essential trait when applying the tool to new or larger areas because vegetation species and composition is variable along most coastlines globally and vegetation species may change due to human actions, such as land clearance (Vijay et al., 2016). The ability of VEdge_Detector to discern a continuous boundary when vegetation was interrupted was less consistent. When vegetation was interrupted by ploughed agricultural land or mangrove clearance a continuous edge was produced (Figure 4.8; Figure 5.16) but gaps in the vegetation edge position were identified when urban land covers interrupted the seaward vegetation boundary (Chapter 5). Current ML and non-ML tools are not able to consistently identify a continuous vegetation edge, where the vegetated land cover is fragmented. Along coastlines containing very fragmented vegetated boundaries, it may be necessary to continue to use manual techniques or use a separate shoreline proxy.

6.1.1.2. Specificity

Alongside tool accuracy, another consideration regarding the performance of a tool is its specificity, whether it exclusively detects the feature of interest or also detects other irrelevant edges. CNNs show promise in being able to discriminate between the coastal vegetation edge and other edges that have similar spectral properties (Chapter 4) whereas all kernel-based operators detected many inland boundaries, not related to the coastline (Figure 3.8 (a) – (h)). This can be attributed to the kernel-based operators identifying the locations with the greatest gradient in pixel values (Liu and Jezek, 2004); these commonly include inland agricultural and urban boundaries. In comparison, the training process provides supervised CNNs with semantic information, enabling the tool to differentiate between the coastal vegetation edge and other vegetation boundaries (Section 4.2.2.3). VEdge_Detector could not discard all inland boundaries, for example between exposed sand and dune vegetation at Wijk-aan-Zee, Netherlands (Figure 4.9 (d)), and inland exposed rock and vegetation in Varela, Guinea-Bissau (Figure 4.6 (d)). This is likely attributed to the similar spectral properties between inland rock and sand, and intertidal zone substrates. Further research should investigate whether the use of a larger training dataset, containing images capturing a more diverse range of coastal locations and features, could enhance the ability of a trained ML tool to exclusively detect shoreline positions from remote sensing imagery.

The ability of a trained CNN to exclusively detect a particular feature, such as the coastal vegetation edge, is an advantage if stakeholders are interested in this specific task but this specificity may hinder the value of the trained tool if the task required is slightly different or broader in scope. CNNs can be proficient at conducting very specific tasks but may less readily transfer to a separate or broader task compared to non-ML tools (Yosinski et al., 2014). For example, a CNN trained to exclusively detect one form of edge is less well suited to identifying the distance between the shoreline and another land cover boundary of interest. In contrast, kernel-based methods and NDVI threshold contours will likely detect these other vegetated boundaries as well, from which the distances between the lines can be calculated. The ability of CNNs to exclusively detect a particular feature is, therefore, only of benefit if this exactly matches the task a stakeholder requires of it.

6.1.2. Can shoreline detection tools cope with Big Data?

The unprecedented rise in publicly available remote sensing imagery poses the question of whether a tool can process, and extract the shoreline position from, such Big Datasets fast enough to be of use in coastal risk management decision making. Google Earth Engine, for example, adds 6000 new images to its catalogue every day and thousands of images of most coastlines globally have been captured (Gorelick et al. 2017). As more imagery is captured, the time series available for every coastal location is lengthening, in turn increasing the length of shorelines which need to be detected in repeated imagery. Where stakeholders are interested in determining shoreline position over a large spatial scale and/or use repeated imagery at fine temporal resolution, it may be important to consider how long it will take for a tool to detect shoreline position from multispectral remote sensing imagery.

The speed of five tools used within this thesis, VEdge_Detector, SVM, NDVI threshold contouring, CoasTool and Canny edge detection, are compared for three sites below. When applied to smaller studies sites (shoreline < 5 km), all tools, except SVM, detected the coastal vegetation edge at a rate of greater than 1 km shoreline length s^{-1}. Along longer, engineering-scale, stretches of coastline, VEdge_Detector speed was almost double the speed of threshold contouring and over five times the speed of Canny-edge detection and CoasTool. SVM was the slowest tool at all three sites, and nearly 50 times slower than VEdge_Detector at Guyana (Figure 6.1). SVM has previously been reported as slower than CNN, with some studies needing to subset an image before applying SVM, because SVM was so computationally expensive (Hasan et al., 2019).

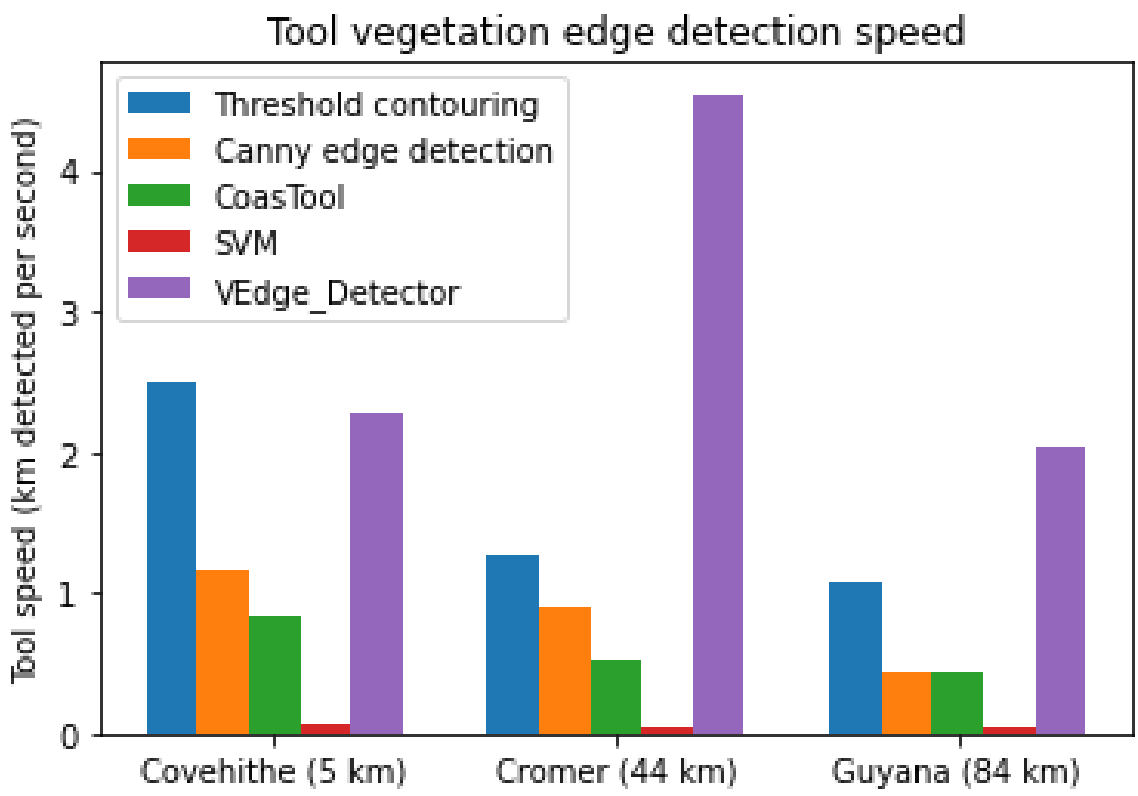


Figure 6.1: Relative speed of five tools used in this thesis to detect the position of the coastal vegetation edge. Speed measured as the number of kilometres of shoreline detected per second. The speed of each tool was compared at three sites of different sizes.

Tool speed is an important consideration for investigations into global rates of shoreline change. This can be demonstrated by considering the contrast between the findings of this thesis and the findings contained within the Intergovernmental Panel on Climate Change (IPCC) 6th report. The IPCC report claims that the region containing the Guiana coastline has accreted by an average of 0.25 m year^{-1} between 1984 and 2015, and that many regions in and adjacent to the Guiana coastline are projected to prograde during the first half of the 21st century (Ranasinghe et al., 2021). This contrasts with the finding in this thesis that 65 % of transects experienced rates of shoreline change greater than ±10 m year^{-1} between 1990 and 2020. The disparities in findings is attributed to the recent IPCC report being based on papers that have estimated global rates of shoreline change using Landsat imagery from just two dates (Mentashi et al., 2018; Luijendijk et al., 2018; Vousdoukas et al., 2020). The Guiana coastline was also determined in this thesis to experience negligible net sediment loss or gain between 1990 and 2020, but experienced rapid rates of erosion and accretion between these times period. The use of images from just two dates, therefore, leads to an underrepresentation of coastal zone dynamics. To ascertain other locations experiencing oscillating or non-consistent patterns of shoreline change at a global scale, future studies require global shorelines to be extracted from images captured on multiple dates, resulting in processing speed being an essential consideration in shoreline detection tool selection.

The importance of tool speed will be dependent upon the length of coastline of interest and the frequency of imagery used, which in turn needs to be compatible with the coastal processes of interest. The extraction of a time series of shoreline position at supra-national scale identified dynamics not detectable using a single image, or imagery with smaller spatial coverage. These dynamics included the identification of locations where coastal receptors were most exposed to shoreline change, identification of spatially separated coastlines with similar shoreline dynamics, and determination of erosional or accretional features migrating cross-shore (chapter 5). High temporal resolution imagery is necessary to, for example, identify changes in shoreline position driven by fluctuations in fluvial sediment supply (Hein et al., 2019), or identify beaches which exhibit a cyclical response pattern to successive storm events (Splinter et al., 2014; Scott et al., 2016). The instantaneous extraction of shoreline position could also aid real-time operations, such as marine navigation and emergency incident response (Yang et al., 2018). Tool speed may be a less important consideration if coastal stakeholders only require an individual snapshot of shoreline position and/or require information on shoreline position along a short stretch of coastline. A low number of images at small spatial scale has, for example, been sufficient for identifying coastal buffer zones where new building developments should be limited (Defra, 2006), and determining localised shoreline response to the implementation of human shoreline modification structures (Hagenaars et al., 2018; Elkafrawy et al., 2021). Even in these scenarios, a timeline of shoreline position may further aid understanding of shoreline response at local scales, thus increasing the need for a high-speed tool.

In summary, tools with fast processing speeds will be important for studies interested in detecting shoreline change at large spatial scales and at fine temporal resolution. As the time series of remote sensing imagery continues to grow, the processing speed will become an increasingly important consideration during tool selection, as more and more images become available for every area of interest. Alongside the speed at which the trained tools can be applied, it may also be important to consider the time and resource required to train and develop the tools in the first instance.

6.1.3. How onerous is it to train and develop a tool?

The time, knowledge and expertise required to train and develop a tool will be an important consideration when an 'off the shelf' tool is not already available to carry out the task required. Whilst in this thesis a trained CNN was faster than other ML and non-ML tools at identifying the position of the coastal vegetation (Figure 6.1), the training and development requirements of CNNs can be very onerous. Here VEdge_Detector required greater than 7 hours to train, compared with the 3 hours 45 minutes needed to train SVM (Chapter 4); and non-ML tools require no training process. In general, the training of CNNs, including the labelling and development of training datasets, can take many weeks and requires the use of high performance computers which are not available to the general public (Gu et al., 2019). The labelling and development of the training dataset in particular is identified here as a bottleneck in future applications of ML to coastal dynamics.

Large training datasets, containing the original and labelled images, are required to train supervised ML tools and more specifically CNNs. The recent explosion in publicly available remote sensing imagery has not been matched by an increase in labelled datasets required to train supervised ML tools (Ma et al., 2019; Tsagkatakis et al., 2019), necessitating the development of these labelled datasets. A dataset of 30,000 labelled paired images was used to train VEdge_Detector and took approximately eight weeks to generate (Rogers et al., 2021). Despite the size of this dataset, it was still only a fraction of the data used to train many ML tools in computer vision studies (Deng et al., 2009). For example, ImageNet, used to detect everyday objects in natural RGB images, contains 14 million labelled images (Deng et al., 2009; www.image-net.org/). Lower VEdge_Detector performance in some coastal settings, for example along rocky cliff coastlines (Figure 4.7) and muddy coasts (Figure 5.4), may be attributed to fewer training images of these coastal features. There is no guarantee, however, that increasing the size of the training dataset will improve performance and the number of images required to train an ML tool to conduct a task will always be problem dependent (Shahinfar et al., 2020). This may be a barrier to the training of future supervised ML tools in the geosciences, due to the risk of a large amount of time and resource being spent on the training of a tool that ultimately exhibits poor performance. To address this issue collaborative work between different academic institutions and citizen science, discussed further below, could be used to generate larger training datasets, potentially analogous in size to those available in computer vision.

Alongside the requirement to manually digitise or label the training datasets, there is also the necessity to develop code using computer programming languages to initially train and develop the different tools. Non-ML tools can be run with no programming knowledge, using software containing a graphical user interface, including ArcGIS and ERDAS Imagine. However, the ability of programming languages to automate tasks and reduce processing time provides a distinct advantage when applying tools to multiple images (Toms, 2015). Despite the complex methods inherent in ML techniques, libraries in programming languages, such as R and Python, enable both ML and non-ML tools to be developed using tens to hundreds of lines of code (Ketkar, 2017). These libraries also enable ML and non-ML tools to be trained and developed, without the need to have a deep understanding of the mathematical principles underpinning the techniques (Gulli and Pal, 2017). Numerous online forums and webpages also provide sample code and step-by-step instructions to develop the different edge detection methods (Sharma, 2017; Brownlee, 2019). The combination of the availability of programming language libraries, and the wide array of online training resources, makes the expertise and time required to write code for ML verses non-ML tools less of a consideration during tool selection.

In summary, the training requirements of supervised ML tools are onerous, particularly the development of labelled datasets, which poses a potential barrier to the future applications of supervised ML tools in the geosciences. Once trained, it has been shown that CNNs are faster than other ML and non-ML tools when applied to new images (section 6.1.2). By combining these two factors, it may be that the training of supervised CNNs is only beneficial when the tool will subsequently be used in many different settings. Alternatively, if a tool is only needed to detect the edges or features from one small-scale image, it is unlikely to be time efficient to train and develop a new supervised ML-tool for this task. In a situation where a trained ML tool has previously been developed and is available for use, tool selection may also consider whether the methods can be explained to a non-technical audience to understand how the tool derived its outputs.

6.1.4. How does it work? Peering into the "black box" of machine learning

The performance of ML tools is enigmatic because it is not clear how they achieve the results they do. CNNs are also commonly considered to be fragile, whereby small perturbations in the spectral properties of an input image can cause large differences in tool output without it being possible to explain the reasons for this change in performance (Ghorbani et al., 2018). The fragility of CNNs can have potentially dangerous consequences, for example applying small stickers to stop signs led to driverless cars interpreting them as speed limit signs (Gu et al., 2019). The fragility and lack of interpretability of ML tools not only affects their performance but also the confidence and trust stakeholders place in them. To overcome this issue, it is necessary to devise tools to shed light into the 'black-box' nature of ML tools. This could improve understanding of why the performance of CNNs varies between imagery and which pixels within each image are important for the CNN when making predictions. To address this issue, interpretability tools have recently been devised that attempt to address the question of how and why the CNN achieved its recorded performance (Bach et al., 2015; McGovern et al., 2019; Toms et al., 2020).

Layer-wise Relevance Propagation (LRP) is a CNN interoperability tool which produces a heatmap identifying the most important pixels or bands, within the input image, used by the CNN to make predictions (Montavon et al., 2018; Toms et al., 2020). After a trained CNN makes a prediction on a new image, LRP can trace how information 'flowed' through the CNN from the input image to the prediction (Barnes et al., 2020). LRP starts at the output node and works backwards through the CNN, determining the nodes with the highest activation values in the hidden layers, and eventually the pixels in the input image which contributed the most towards the NN prediction (Bach et al., 2015; Montavon et al., 2018). LRP has been used to produce a heatmap of the most important pixels for a CNN trained to locate convection clouds and storms in multispectral imagery (Ebert-Uphoff and Hilborn, 2020; Hilborn et al., 2020; Lee et al., 2020). Applying these techniques to CNNs trained to detect shoreline position may aid understanding of why the tool's performance varies in time and location. Thus, for example, whether or not different intertidal substrates or tide heights affects CNN performance.

Permute-and-predict (PAP) is another method which has been applied to increase the interpretability of NN outputs (Jergensen et al., 2020). After the trained NN has made initial predictions, the values of one or more of the input parameters are permuted and the impact this has on the output predictions are calculated (Barnes et al., 2020). The most important parameters are inferred to be those which make the greatest changes to NN predictions when their values are altered (McGovern et al., 2019). Andersson et al. (2021) predicted future sea ice extent by training a NN on time series of previous sea ice extent and climate parameters. The parameters which influenced the performance of the NN the most were compared to pre-existing domain knowledge of the causes of changes to sea-ice extent. This provided greater confidence that the NN was identifying meaningful relationships between different conditions (Andersson et al., 2021). Both PAP and LRP are examples of interoperability tools that could improve understanding of the layers and pixels most important for the ML tool when making predictions. This in turn could provide greater reassurance that the CNN is using information consistent with local knowledge on the most important factors in determining coastal risk.

In flood and coastal risk management (FCRM), model interpretability is a key consideration for local communities when accepting the outcome of the tool as justification for the implementation of a flood or coastal risk mitigation or adaptation measure (Maskrey et al., 2019). Local stakeholders, including homeowners and business owners, may render a tool untrustworthy if it cannot be explained how the results were derived. Alternatively, if a tool is not understood, the inherent limitations and uncertainties within the methods may not be recognised, potentially leading to overconfidence in the use of the model (Voinov and Gaddis, 2008; Maskrey et al., 2019). In many countries, such as the UK, the ability of stakeholders to understand how models generate their outputs is increasingly essential where a participatory approach to risk management is being employed, whereby local communities have a greater influence on the risk management decision making process (Environment Agency, 2020b). Where it is not possible to explain the outputs of a tool to local communities, there can be an increasing feeling that a 'top-down' approach has been taken, whereby people without local knowledge make decisions, whilst disregarding the views of people who live or reside in the area (Voinov and Gaddis, 2008). For ML tools to be further employed in flood and coastal risk management schemes, tool interpretability will become an essential consideration.

In summary, the inner workings of CNNs remain elusive. In comparison, it is possible to, for example, manually study the NDVI layer of a coastal scene to understand why threshold contouring did or didn't work, and whether a different threshold value may improve results. Interoperability tools, including LRP and PAP could be used to aid understanding of how and why CNNs produce their outputs. The code for interpretability methods is increasingly being made publically available on platforms such as Github (Alber et al., 2019; Li et al., 2019; Toms et al., 2020), and future research applying CNNs in the geosciences should concurrently publish the results of interpretability methods such as those described above. For ML specialists, this could aid decision making on changes which need to be made to the training dataset or CNN architecture to improve performance. For the wider research community and coastal stakeholders, greater interpretability may increase the trust placed in CNNs as a set of tools to complement pre-existing knowledge and increase the likelihood of future collaborations between the two sets of research groups.

6.1.5. What resolution imagery can the tools be used on?

The spatial resolution of remote sensing imagery is a key determinant of the coastline dynamics that can be monitorred. Landsat, Planet, and vertical aerial imagery with 30 m, 3 – 5 m and sub-metre spatial resolution respectively constituent some of the main forms of publicly available multispectral remote sensing imagery. Differences in their spatial resolution mean they can each detect different coastal dynamics. Determining whether a particular tool can be applied to multispectral imagery with different resolutions is necessary for stakeholders interested in determining different scale coastal processes. No study has ever used Planet's 3 m resolution PlanetScope imagery to detect shoreline change and only one study has used 5 m RapidEye imagery for waterline change detection (DaSilva et al., 2021). Thus, it can be concluded that the performance of different edge detection tools when applied to different types of remote sensing imagery remains undetermined.

Landsat imagery is the conventional source of remote sensing imagery applied in shoreline change studies (Toure et al., 2019). It has been demonstrated that ML and non-ML techniques can be applied to Landsat imagery to discern shoreline dynamics at a multitude of scales, including site-based (Liu and Jezek, 2004; Kuleli et al., 2011; Nassar et al., 2018; Pardo-Pascual et al., 2018), supra-national (Chapter 5) and global-scale (Pekel et al., 2016; Luijendijk et al., 2018; Mentashi et al., 2018). In other circumstances, finer resolution imagery is necessary to detect annual to decadal shoreline change. In this thesis, the greatest amount of retreat at Covehithe, Suffolk, over the 11 year study period was 54.38 m (Chapter 4). Annual-scale shoreline dynamics, and shoreline response to major storm events, were detectable using Planet imagery due to its fine (3 – 5 m) spatial resolution (Figure 4.10). However, 54.38 m measures less than two Landsat pixel widths, precluding the use of this imagery to detect decadal-scale dynamics at Covehithe. High magnitude, low frequency, storm events are key drivers of shoreline change in many coastal zones (Brooks and Spencer, 2012; Masselink and Van Heteren, 2014; Brooks et al., 2017) making it essential for future studies to use remote sensing imagery, such as Planet, which can discern dynamics at this scale. Even along dynamic stretches of shoreline, few locations will experience annual rates of shoreline change greater than ±60 m, rendering Landsat imagery unsuitable for detecting this temporal scale of change. Despite the superior spatial resolution of Planet imagery enabling the detection of shoreline dynamics not discernible using Landsat imagery, the shorter, 11-year, temporal coverage of Planet imagery means it is still necessary to apply edge detection tools to Landsat imagery when investigating multi-decadal shoreline change.

In this thesis, CNNs were able to identify ~90% of the coastal vegetation edge in Guiana (Chapter 5), even though no Landsat imagery was contained within the training dataset (Chapter 4). Conversely, VEdge_Detector was unable to extract a vegetation edge from 50 cm resolution aerial imagery (Chapter 4). This is attributed to the greater intra-class variability in vegetation species spectral properties when using finer resolution imagery (Hasan et al., 2019; Liu et al., 2020). The findings here suggest that a CNN trained using imagery of a particular spatial resolution may be able to detect features and edges in coarser imagery but not necessarily in finer resolution imagery. Whether it is a limitation that CNNs cannot be applied to finer resolution imagery depends upon the scale of the coastal dynamic of interest. Processes which are not detectable using 3 – 5 m resolution imagery are unlikely to increase the risk of erosion to a coastal receptor. Alternatively, if, for example, a cliff top property would erode into the sea in a scenario where there was less than 3 – 5 m of erosion, the property owner and other coastal stakeholders are unlikely to require a remote sensing-based study to inform them of this fact.

In summary, with the current availability of satellite imagery, Landsat imagery remains a valuable resource for identifying regional, decadal-scale shoreline change hotspot locations. However, finer resolution imagery, such as Planet, should then be used to detect annual-scale dynamics along shorter stretches of shoreline. From a management perspective, the scale of shoreline change dynamics of interest will be the key determinant of the type of imagery used. As demonstrated, CNNs trained to detect shoreline position are currently able to detect shoreline position in imagery with 3 m resolution and coarser. Shoreline dynamics which are not detectable using 3 m resolution imagery are, however, unlikely to generate large changes in erosional risk to coastal receptors. As new satellite platforms continue to be launched, including those collecting 50 cm resolution Planet imagery (Fu et al., 2020; Masek et al., 2020), the selection of imagery used to train future ML tools should consider whether the dynamics of interest can be discerned using this imagery.

6.1.6 So is ML the way forward?

The choice of tool used to detect the coastal vegetation edge will be problem, scale and location specific. Along shorter stretches of shoreline, most ML and non-ML tools have similar accuracy and processing speeds. In these circumstances, the current inability to readily interpret the inner workings of ML tools, especially CNNs, combined with their large training and development requirements, may become the decisive factors during tool selection. Due to these limitations, local stakeholders may see ML tools as less appealing compared with more established, non-ML techniques, including threshold contouring and kernel-based edge detection operators. Along supra-national to global scale stretches of coastline, the ability of ML tools to detect the coastal vegetation edge under different biotic and abiotic conditions, combined with the superior processing speeds, may become more influential factors. The ability of ML tools to generalise also opens opportunities for them to be used in other coastal zoneswhich have not previously benefitted from detailed studies. Irrespective of the relative performance and other attributes of ML tools in different scenarios, they will only be selected if third parties have trust in their output. This trust will only develop if methods to improve ML tool transparency in operations and performance are generated.

Prior to the establishment and widespread application of these interpretability methods, the simultaneous use of multiple vegetation edge detection tools may be of benefit in reassuring stakeholders of the validity of ML and non-ML tool outputs. If multiple tools are consistently identifying the position of the vegetation edge, and associated rates of change, this may provide greater reassurance to third parties that the outputs are robust. Where inconsistencies in outputs are determined, this would also be of great use in identifying scenarios and locations where different tools perform better or worse. As new remote sensing imagery becomes available, for example from Landsat 9 and 50 cm resolution Planet imagery (Houborg and McCabe, 2018; European Space Agency, 2021; Masek et al., 2020), simultaneous use of multiple tools will also provide information on relative tool performance when applied to these different image datasets.

6.2. Different shoreline proxies, different coastal dynamics

Each shoreline proxy carries different information regarding shoreline change (Boak and Turner, 2005; Toure et al. 2019). These proxies also contain limitations, most notably that their position may change even when there has been no net shoreline erosion or accretion. Information on the coastal dynamics which can be gleaned from using a particular shoreline proxy may be of benefit during stakeholder engagement to decide on the shoreline detection tool(s) to use. This decision will be locally specific and depend upon the rates of shoreline change and coastal receptors present in the area of interest. Previous remote-sensing based studies have almost exclusively focused on waterline detection (Toure et al., 2019) but future studies will benefit from the availability of publicly available tools which can detect multiple different proxies. It is argued here that, despite previous overdependence on the detection of the waterline, along many shorelines changes in the coastal vegetation edge potentially provide better representation of changes to coastal risk.

Along storm-driven coastlines, the coastal vegetation edge has been shown to be more informative than the instantaneous waterline in representing shoreline dynamics (Chapter 4). At Covehithe, the vegetation line remained stable or retreated between every time period (Figure 4.10; Figure 4.13), corresponding with the findings of other studies using ground-referenced measurements or other techniques that have reported consistent retreat along these soft rock cliffs (Brooks and Spencer, 2012; Burningham and French, 2017; Environment Agency, 2010). In comparison, waterline here clearly fluctuates under diurnal and semi-diurnal tides (Figure 4.14), masking the dynamics of consistent shoreline retreat.

This highlights that in many locations globally, the over-reliance on the use of the waterline as a shoreline proxy in isolation may lead to incorrect conclusions being drawn about coastal dynamics, including determining that some coastlines accrete between storm events. The use of the waterline in isolation may, in turn, lead to an under or overestimation in the likelihood of future shoreline retreat events.

The vegetation line has also been shown to be a more robust proxy of shoreline position along a shallow gradient muddy shoreline (Chapter 5). Along this coastline, unrealistic, highly sinuous, waterlines were produced when using the widely applied NDWI threshold contouring algorithm (Chapter 5; Vos et al., 2019b). This resulted in manually digitised shorelines needing to be produced for the VEdge_Detector moving-window algorithm to 'walk' down (Chapter 5). Inaccurate waterline detection is attributed to the similar spectral properties of shallow and turbid water, as well as exposed mud banks (Bishop-Taylor et al., 2019). Current methods to detect waterlines along shallow gradient and muddy shorelines have reported sub-pixel accuracy (Zhao et al., 2008), but these methods required concurrently collected DEMs of the intertidal zone, restricting their use in other coastal zones. In contrast, the ability of mangrove vegetation to rapidly establish on newly formed mud banks enables the coastal vegetation edge to provide a robust representation of shoreline accretion and erosion (Fromard et al. 2004; Gratiot et al. 2008; Walcker et al. 2015). This highlights how the instantaneous waterline is not a panacea proxy of shoreline dynamics and that the vegetation line provides a promising alternative to representing these dynamics along storm-dominated, muddy, and sandy shorelines.

Limitations when using the coastal vegetation edge include changes to the vegetation line position, caused by processes not related to coastal dynamics, including anthropogenic land clearance and seasonal dynamics. Furthermore, multiple tools do not produce continuous vegetation edges along coastlines where the vegetation line is heavily fragmented, for example along urbanised stretches of coastline and where ploughed agricultural fields interrupt the vegetation line (Chapter 3; Chapter 5). Some coastlines also contain forms of vegetation which it is misleading to detect because their movement may be due to processes unrelated to coastal erosion and accretion. This includes the presence of macroalgae at the base of cliffs (Chapter 4). The same issue could arise from the presence of, for example, algal blooms on the surface of tidal flats or on the surface of the water column itself. In locations where these forms of vegetation are present, and the tidal-driven fluctuations in waterline are negligible due to the presence of steep topography or small tidal ranges, the instantaneous waterline may provide a better proxy of shoreline position and dynamics.

The stability of the coastal vegetation edge lends itself to representing sandy beach cliffs (Figure 4.10) and mudflat (Figure 5.11) coastal morphological response to storm events, as has been identified for gravel barrier islands (Pollard et al., 2020). Conversely, however, this stability precludes using vegetation positional change to discern foreshore sediment dynamics. These foreshore dynamics, detectable using the waterline, include sandbar migration and post-storm sediment deposition (Aagaard et al., 2013; Burvingt et al., 2016; Goldstein et al., 2019). These processes are generally smaller than engineering scale dynamics. This thesis supports other studies that have found that vegetation position can show the response of shorelines to individual storm events and other sub-annual processes (Barrett-Mold et al., 2010; Bullard et al., 2019), and the impact of storms on backshore dynamics (Grzegorzewski et al., 2011; Toure et al., 2019) which cannot be detected by using the waterline in isolation. The waterline may provide a better representation of daily to sub-annual coastal sediment dynamicsbut the vegetation line can act as a low-pass filter, highlighting locations of sustained shoreline erosion or recovery. The different rates of change of the two shoreline proxies highlights the benefits which may derive from their simultaneous extraction. Going forwards, future work should investigate which proxies are better suited to different coastal morphologies and temporal scales of shoreline change, to enable future research to make informed decisions about which abovementioned publicly available tool to use to extract shoreline positions.

To aid these investigations, tools for detecting the coastal vegetation edge (github.com/MartinSJRogers), the instantaneous waterline (Dai et al., 2019; Vos et al., 2019b), cliff top position (Payo et al., 2018) alongside others, are now publicly available. This increase in access provides opportunities for the tools to be applied to other locations. With the provision of publicly available, automated tools to extract different shoreline proxies, future research can analyse the proxy(s) that best represent coastal processes of interest, instead of being limited to selecting the proxy that is methodologically possible to extract. All abovementioned tools come with instructions on how they are applied, providing greater equity because they can be used by anyone with suitable expertise. The ability of these tools to be used by local stakeholders may reduce the feeling that a top-down approach has been applied, where individuals with no local knowledge have conducted research into coastal change.

In summary, the availability of multiple automated tools enables the shoreline proxy which best represents the coastal dynamics of interest to be extracted, or even for multiple proxies to be simultaneously extracted from remote sensing imagery. The simultaneous extraction of the water and vegetation line could enable the detection of other features, including beach or intertidal zone width to be identified, which is not possible using one shoreline proxy in isolation. It allows for a more robust indication of net shoreline movement, because both proxies contain limitations, but if both proxies migrate landwards, there is greater certainty that it is due to coastal erosion, instead of tidal influence on water and vegetation land clearance. Simultaneous extraction of multiple shoreline proxies can provide better insight into cross-shore and alongshore dynamics, for example determining gravel barrier response to storm events (Pollard et al., 2020), and changes in beach volume due to sediment nourishment schemes (Wilson et al., 2015; Wilson et al., 2019). Simultaneously extracting multiple shoreline proxies can increase individual tool performance; for example, by extracting the marsh scarp as a shoreline proxy, salt marsh vegetation extent was more accurately digitised. This may have been because marsh vegetation normally ends in close proximity to a distinct drop in elevation (Wang, 2009; Farris et al., 2019). Further investigation could determine whether simultaneously extracted proxies may help to determine other coastal dynamics, which are not detectable using one proxy in isolation, such as a change in beach width or slope in response to storm events.

6.3. Perceiving coastal risk through the machine learning lens

Irrespective of the shoreline detection tool(s) applied, to determine the level of coastal receptor exposure or risk to shoreline change, the outputs of shoreline change tools require integration with information pertaining to other aspects of coastal risk. ML is providing new opportunities to determine rates of shoreline change alongside other factors relating to coastal risk. Most previous studies have used ML to investigate one element of coastal risk in isolation (Goldstein et al., 2019). Discussed are two main advantages of using ML tools to derive risk indices over other methods and alongside other tools and diverse datasets: (i) the ability to rapidly detect and update changes to risk indices at fine spatio-temporal resolution, via the integration of important risk-based data; and (ii) the ability to determine populations exposed to shoreline change at a supra-national scale.

ML tools have commonly been used to provide information on one aspect of coastal risk in isolation VEdge_Detector and WorldPop, for example, independently provide information regarding rates of shoreline change or provide information on population dynamics in the coastal zone respectively. But taken together, their outputs provide information on the relative number of people exposed to shoreline change. A key advantage of using these ML tools is the ability to frequently update the 'population at risk' values. Once trained, ML tools have been shown to be faster than other non-ML tools at identifying shoreline position (Figure 6.1), and the remote sensing datasets used by the ML tools are collected at sub-annual scale (Stevens et al; 2015; Llyod et al., 2019). This enables populations exposed to shoreline change to be calculated at a sub-annual scale. The consistent, gridded nature of many ML outputs enables them to be readily harmonised and integrated, as demonstrated by the ability to calculate relative populations exposed to coastline change at 100 m transect intervals (Figure 5.10), commonly not possible using aggregated census data (Stevens et al., 2015; Tatem et al., 2017). The consistently gridded nature of population datasets obtainable using ML tools also avoids issues such as the modifiable areal unit problem, whereby the choice of boundary condition can heavily affect the representation of population dynamics present (Tatem et al., 2017). The high processing speeds of ML tools, combined with their use of datasets with global spatial coverage, also provides potential in consistently determining coastal dynamics, including in locations which have not previously benefitted from detailed study.

ML tools trained using datasets with global coverage are country-agnostic, meaning they can be consistently applied to detect coastal dynamics in any location globally. This will increase the ability to determine locations where populations are at greater risk to shoreline change as those experiencing rapid erosion and/or urbanisation, instead of locations that have been most heavily researched. A low number of shorelines around the world are extensively researched, including Narrabeen Beach, Australia (Beuzen et al., 2018; Splinter et al., 2018), Covehithe, UK (Brooks and Spencer, 2014; Burningham and French, 2017; Rogers et al., 2021), and Fire Island, USA (Wilson et al., 2015). Research in these locations has advanced understanding of shoreline response to extraneous forcing factors, but the locations benefitting from the greatest amount of research do not necessarily correspond to locations where populations are at greatest risk to shoreline change. Big Data is increasingly capturing many physical and socio-economic processes in the coastal zone (Rumson et al., 2018). This means that it is feasible for future datasets pertaining to vulnerability indices, meteorological conditions, and hydrological conditions, to be large and detailed enough to train ML tools to generate thematic layers of different aspects of risk. ML tools which utilise datasets of global coverage provide promise in generating consistent levels of information on levels of exposure and risk globally, including in currently understudied locations. There is a need, however, to ensure equity in this knowledge production, for example ensuring that local communities or stakeholders are not disadvantaged or excluded from the decision-making process if they do not have access to computer resources or Big Data facilities.

ML tools individually applied to different aspects of coastal risk could be integrated to, for example, improve understanding of hazards derived from wave action in the coastal zone. ML tools have been applied to synthetic aperture radar (SAR) data, a separate form of remote sensing imagery, to automatically identify wave locations and wave conditions, including wave height, wave incident angle and return period (Wang et al., 2019; Wu et al., 2021). These ML-derived datasets provide more consistent coverage of wave conditions than is possible using wave buoys or other forms of ground-referenced measurements (Collard et al., 2009). Separately, other ML tools have been used to automatically identify the position of hurricanes and other coastal tropical storms in remote sensing imagery (Kim et al., 2019; Lee et al., 2020), proving regularly updated information on storm locations and severity at similar spatial resolution to the wave data. The integration of these ML datasets could aid understanding of air-sea interactions in coastal zones, in particular wave response to different meteorological conditions, or determine how waves build up in different locations when subject to similar storm events (Topouzelis and Kitsiou, 2015). This could, in turn, identify wave height and energy hotspot locations, where wave condition are consistently greater and adjacent coastal receptors are subsequently potentially more exposed to wave action and shoreline retreat.

The outputs of the abovementioned ML tools investigating coastal hazards could further be integrated with ML tools investigating the exposure and vulnerability of receptors in the coastal zone. Alongside ML tools to estimate population density, ML tools have been used to classify land cover classes and associated rates of urbanisation in coastal regions (Rogan et al., 2008; Karpatne et al., 2016). ML tools have also determined near-instantaneous damage to buildings and communities caused by hurricanes and other tropical storms (McCarthy et al., 2020; Yuan and Liu, 2020). Integrating ML tools investigating damage to communities and buildings with those determining real time information on tropical storm position, may be of real benefit to incident response units, to determine current or predict the imminent impacts caused by the storm. More broadly, the production of multiple ML-derived datasets pertaining to different risk indices could be integrated to inform risk management decisions and support policy making in the coastal zone. This includes policy to restrict rates of urbanisation and infrastructure development in areas experiencing rapid rates of shoreline change, or promote the protection or management of coastal ecosystems to provide further protection to adjacent land covers. These policies would both limit the number of people vulnerable to future coastal recession, as well as reducing degradation to the health of coastal ecosystems and resources (Neumann et al., 2015). However, ML tools are unlikely to be able to be trained on all aspects of risk. Thus, for example, large datasets pertaining to stakeholder perception of the most important receptors to protect are not currently available, primarily because stakeholder views vary substantially between and within different regions (Penning-Rowsell et al., 2014). ML remains a tool to inform, instead of replace, stakeholder-led risk-management decision making.

Other limitations of ML tools, particularly those exclusively using remote sensing-based data, must also be considered when choosing whether to apply the tools to coastal risk studies. The use of remote sensing data precludes the detection of some smaller scale dynamics pertaining to coastal risk, for example, WorldPop data did not detect some smaller settlements, such as Shell Beach, Guyana (Chapter 5), emphasising the need for remote sensing data to be complemented with local knowledge and data gathering. This is attributed to ML tools used remote-sensing derived ancillary data of Earth features like ‘night light’ and ‘road density’, which may not provide evidence of smaller settlements in rural locations (Tatem et al., 2017). Remote sensing imagery, commonly used in ML applications to coastal risk, cannot directly measure all aspects of coastal risk, meaning some information has to be inferred. ML applications can be used to, for example, identify hurricane and convection cloud locations, but information on the severity of the storm, including wind conditions, can only be estimated (Lee et al., 2020). This highlights that ground-referenced studies are still required to validate the performance of ML tools and identify dynamics not detectable using ML in isolation and that ML tools should not replace the need for censuses and other ground-based data collection methods.

In summary, ML tools can be used to consistently determine population dynamics and shoreline change over large spatial areas and fine resolutions. Further research should investigate their potential to be applied to other risk indices but caution should be taken to ensure they do not replace, but instead complement, stakeholder engagement in coastal risk management decision making.

6.4. Concluding remarks

6.4.1. Machine learning in the geosciences: future potential and collaborations

The rise of Big Data and ML techniques has motivated many investigations into whether or not ML can advance knowledge in different research domains. Quantifying the benefits, or otherwise, provided by ML techniques is best achieved by comparing the performance of ML tools with established, non-ML, methods. To this end, VEdge_Detector, a convolutional neural network (CNN), was developed for the automatic detection of the position of the coastal vegetation edge in multispectral remote sensing imagery. Its performance was compared to a range of ML and non-ML techniques. VEdge_Detector shows promise in generalising to detect the vegetation edge in a range of coastal settings at local to supra-national scales, whilst producing similar or smaller positional errors compared with other methods (Chapter 3, Chapter 4). From a coastal stakeholder perspective, tool accuracy is not the only consideration when selecting a method to identify edges and features in remote sensing imagery. A tool will be chosen depending upon its processing speeds, training and development requirements and interpretability. CNNs have been shown to be faster than other edge detection tools (Figure 6.1). However, a potential barrier to the further uptake of supervised ML techniques is the onerous nature of developing reference datasets. Despite the training process providing VEdge_Detector with semantic information to distinguish between different edges and features in the imagery, there were still limitations in its performance, for example along rocky and urbanised stretches of coastline. Increasing the size and diversity of the imagery in the training dataset may improve the performance of the CNN but moving forwards, the training of future CNNs should investigate the use of collaborative research and citizen science.

One way that large, labelled datasets could be generated is through greater collaboration between disparate academic research groups or citizen scientists interested in detecting the same features in remote sensing imagery. Researchers across different research fields including earthquakes, volcanoes, wildfires and flood risk have all contributed to labelling a dataset of damaged/undamaged buildings in remote sensing imagery (Gupta et al., 2019). This dataset subsequently trained a CNN to detect building damage caused by coastal storms (Chen et al., 2021), highlighting the transferability of the dataset. These examples highlight that, irrespective of the natural disaster itself, participating research groups can work together to develop effective datasets that can be used to train ML tools for a range of tasks. This collaborative approach could also be applied to detect vegetated features in remote sensing imagery.

The detection of vegetated features and edges in remote sensing imagery can be important for research in the fields of plant health (Hamdi et al., 2019; Sylvain et al., 2019), coastal risk management (Rogers et al., 2021), land cover change (Buscombe and Ritchie, 2017), and agricultural crop type (Hoeser et al., 2020). An example of a dataset, derived through collaboration, which could be repurposed to explore an alternative means includes that developed by Weinstein et al., (2021), an augmented labelled dataset of over 100 million tree crowns. This dataset was originally produced to analyse tree health but could be of benefit to detect forest clearance, expansion of urban extent and plant-climate response. Citizen science could additionally contribute via tools such as ‘Make sense AI’ (https://www.makesense.ai/), which allow members of the public to label remote sensing imagery (Buscombe and Ritchie, 2017). The use of these collaborative or citizen science approaches could both reduce the burden of developing a dataset, and also help to provide labelled data which could be applied to the training of multiple ML tools investigating different tasks.

Research groups could also collaborate to produce a centralised, large-scale repository of multiple different tools that utilise ML-based techniques to extract features from remote sensing imagery. One early example is SciVision, which aims to become a repository or ‘search-engine’ for different ML-based tools produced by different research groups globally (https://github.com/alan-turing-institute/scivision). If enough tools are added to a repository such as SciVision, it would then be possible for third-parties to search for tools by particular criteria e.g., remote sensing imagery type, image spatial resolution or feature of interest. A repository of this nature would enable other researchers to utilise or further develop the code produced within these tools. This in turn may reduce duplication caused by different research groups generating very similar tools. Further, a repository of this nature may increase the exposure or likelihood of third-party non-technical stakeholders discovering and utilising the tools for various purposes. Repositories like SciVision are necessary because even when a ML-based tool is made publicly accessible on platforms such as Github, they will only be discovered if people know to look on that particular repository page. A central repository may aid in the wider dissemination and utilisation of ML-based tools that extract features and knowledge from remote sensing imagery.

6.4.2. Making the most of the rise in Big Data

Alongside applying ML tools to detect edges and features in multispectral remote sensing imagery, their outputs provide promise in accurately representing rates of shoreline change. In this thesis, VEdge_Detector outputs were used to identify rates of shoreline change in a range of locations globally, focussing on Covehithe, UK and the Guiana coastline, northern South America. Supra-national scale shoreline detection enabled the identification of coastal processes not detectable using smaller-scale imagery, including migrating erosion hotspots, as well as spatially distinct stretches of coastline exist which portray very similar oscillations in erosion and accretion (Chapter 5). These findings have been possible due to the marked increase in publicly available remote sensing imagery. Future developments in the spatio-temporal resolution of multispectral imagery could provide further insights into coastal dynamics, not currently discernible using the imagery available.

Planet's mission to "Scan the whole earth every day" with its flock of satellites provides promise in the collation of fine resolution imagery over a large spatial area (Planet Team, 2017). For example, between January and August 2021, 252 and 355 PlanetScope images captured the entire study areas of Shell Beach (Guyana) and Covehithe (UK) respectively, equating to between eight and 11 images per week (Planet API, 2021). The use of many satellite platforms concurrently capturing imagery of the Earth's surface will reduce dependence upon the operational performance of a low number of sensors. Along many coastlines globally, less than 100 Landsat images are available for each scene (Pekel et al., 2016). The number of images which can be analysed is further reduced by cloud cover and severe stripping on Landsat 7 imagery (2003 onwards), caused by damage to the Scan Line Corrector (Young et al., 2017). This resulted in concurrent imagery for the entire Guiana coastline being available during only six years between 1990 and 2020 (Section 5.2). The large number of Planet satellites could reduce dependency upon the performance of an individual satellite and image capturing device and limit the impact caused by the presence of clouds in one image. This will hopefully increase the temporal density of imagery available from which to extract shoreline position.

Advancements in the temporal resolution of both Planet and Landsat will also enable imagery to be captured at more similar times each year, imperative when studying changes to vegetation extent or boundaries. When calculating rates of shoreline change at both Covehithe and Guiana, it was not possible to collect imagery at the same time every year and concessions had to be made so that all imagery used was captured within the same four-month period of that respective year. Imagery was selected to correspond to maximum summer vegetation coverage in the UK and the dry season, with lower cloud cover, in Guiana (Ballère et al., 2021; chapter 4, chapter 5). Collecting imagery at the same time of year is particularly pertinent in vegetation edge studies because salt marsh (Möller and Spencer, 2002; Möller, 2006), dune and mangrove (Jana et al., 2016) vegetation coverage can vary substantially between seasons, irrespective of coastal zone dynamics. Advances in the temporal resolution of remote sensing imagery will provide opportunities to collect imagery from the same time of year to enable more robust comparisons of annual-scale change in coastal vegetation extent.

The degree to which the geosciences can benefit from this increase in remote sensing imagery is dependent upon the costs of attaining the data. An academic licence permits free access to 10,000 km^2 of Planet imagery per month (Marta, 2018). However, the eight Landsat scenes encompassing the Guiana study site cover an area greater than 250,000 km^2 (Figure 5.1), meaning that many years would be required to download sufficient Planet imagery. To gain access to more imagery, most satellite companies require academic institutions to acquire licences for huge amounts of imagery (Marta, 2018). This is costly and only a fraction of the data allowance is likely to be used. These financial restrictions currently preclude the wide-scale application of Planet imagery into supra-national scale studies. Removing the paywall separating the imagery from the research community could provide greater equity, whereby it is not only academic institutions in wealthier countries that are able to conduct this research. To cover the costs of launching the satellites and capturing the imagery, whilst allowing more equitable access to the imagery, it is suggested that satellite companies could charge per unit of data acquired, instead of using the current licencing approach.

6.4.3. Predicting our future coast

The temporal coverage of remote sensing imagery continues to increase, providing an opportunity to use edge detection tools to generate a time-series of historical shoreline positions which in turn could be used to train a separate ML tool to predict future shoreline position. Other studies using Bayesian Networks have predicted shoreline change using a time-series of historic shoreline position, combined with datasets on wave conditions (Beuzen et al., 2018; Giardino et al., 2019). These studies show promise in using ML tools to predict shoreline change. However, they exclusively used ground-referenced measurements of shoreline positionwhich are not available in most locations. In 20-30 years, multi-decadal time series of fine resolution imagery, including Planet's 50 cm resolution products (European Space Agency, 2021), will provide opportunity to train ML tools on sub-annual scale processes not currently detectable using Landsat or other coarser imagery. To make robust predictions of future shoreline position, datasets pertaining to external forcing factors are also required.

Meteorological and hydrological data with sufficient spatio-temporal coverage and resolution is necessary to train an ML tool to predict future shoreline positions. Tide and wave gauges, providing observation-based data, are unevenly distributed, restricting their ability to detect spatially heterogeneous wave conditions, even in relatively data rich regions like the UK (Rumson and Hallett, 2018; National Oceanography Centre, 2021b). Lessons could be learnt from the field of climate science where ML tools used to predict future climatic conditions are commonly trained on synthetic datasets generated by process-based models and calibrated using observational data (Rasp et al., 2021). Compared with observational datasets, process-based models can generate synthetic datasets with longer time-series and finer spatial resolution (Dueben and Bauer, 2018; Gentine et al., 2018). Synthetic datasets of rare storm events can also be generated, key drivers of shoreline retreat in some coastal zonesbut for which very few observed data entries exist (Dueben and Bauer, 2018). Wave models with large spatial coverage and finer and gridded resolution, for example ‘WaveWatch III’ (National Oceanic and Atmospheric Administration, 2021) could be used to train a ML tool on historic oceanic conditions. Using synthetic data from process-based models enables physics-based domain knowledge to be incorporated into ML training (Kashinash et al., 2021). However, ML tool performance is currently limited by the capabilities of the underlying process-models (Dueben and Bauer, 2018).

Human shoreline modification factors are another key determinant of human shoreline position. There is scope to train a supervised ML tool to automatically detect human shoreline modification structures via multispectral imagery. Referenced datasets already exist, as detailed inventories of the human shoreline intervention type and position are maintained in countries such as the UK and The Netherlands (Rumson and Hallett, 2018; Environment Agency, 2021). Many studies have used remote sensing imagery to identify or predict shoreline response to sea defence installation (Giardino et al., 2019; Elkafrawy et al., 2021) but these commonly utilise local knowledge, or locally accessible datasets, to ascertain the location and dates of shoreline interventions. Automatic detection of these features could be particularly valuable in locations where local knowledge or analogous data are not available and could be applied to historical images to determine when defences were installed. These could aid our understanding of the relative influence of human shoreline modification structures on shoreline change drivers of shoreline change, thus informing on how they may influence their future position.

Tools that predict future shoreline position would be central to answering the ultimate question of which locations and coastal receptors are likely to be exposed to the greatest risk to shoreline change in the future. This could be achieved by combining predictions on rates of shoreline change with ML or process-model based predictions of location specific future rates of sea-level rise (Oppenheimer et al., 2019) as well as changes to human populations densities and vulnerability metrics in the coastal zone (Neumann et al., 2015). This information would be very informative for local to supra-national scale planning in the coastal zone, including identifying locations where no further developments should take place. There would also be the potential to produce a ‘timeline’ of when different coastal locations are likely to require the implementation of adaptation or mitigation measures to reduce the impacts of future shoreline change events such as managed realignment and planned inland human migration.

References

Aagaard, T., Greenwood, B. and Hughes, M. (2013). Sediment transport on dissipative, intermediate and reflective beaches. *Earth-Science Reviews,* 124(1): 32-50.

Al-Amri, S.S., Kalyankar, N.V. and Khamitkar, S.D. (2010). Image segmentation by using edge detection. *International Journal on Computer Science and Engineering*, 2(3): 804-807.

Alber, M., Lapuschkin, S., Seegerer, P., Hägele, M., Schütt, K.T., Montavon, G., Samek, W., Müller, K.-R., Dähne, S. and Kindermans, P.J. (2019). iNNvestigate Neural Networks! *Journal of Machine Learning Research,* 20(93): 1–8.

Allen, J.R.L. (2000). Morphodynamics of Holocene salt marshes: a review sketch from the Atlantic and Southern North Sea coasts of Europe. *Quaternary Science Reviews,* 19(12): 1155–1231.

Allison, M.A. and Lee, M.T. (2004). Sediment exchange between Amazon mud banks and shore-fringing mangroves in French Guiana. *Marine Geology*, 208(2-4): 169–190.

Allison, M.A. and Lee, M.T., Ogston, A.S., Aller, R.C. (2000). Origin of Amazon mud banks along the northeastern coast of South America. *Marine Geology,* 163(1): 241–256.

Almonacid-Caballer, J., Sánchez-García, E., Pardo-Pascual, J.E., Balaguer-Beser, A.A. and Palomar-Vázquez, J. (2016). Evaluation of annual mean shoreline position deduced from Landsat imagery as a mid-term coastal evolution indicator. *Marine Geology*, 372(1): 79–88.

Andersson, T.R., Hosking, J.S., Pérez-Ortiz, M., Paige, B., Elliott, A., Russell, C., Law, S., Jones, D.C., Wilkinson, J., Phillips, T., Tietsche, S., Sarojini, B., Blanchard-Wrigglesworth, E., Aksenov, Y., Downie, R. and Shuckburgh, E. (2021). Seasonal Arctic sea ice forecasting with probabilistic deep learning. *Nature Communications,* 5124(1): 1-12.

Anthony, E.J., Brunier, G., Gardel, A. and Hiwat, M. (2019). Chenier Morphodynamics on the Amazon-Influenced Coast of Suriname, South America: Implications for Beach Ecosystem Services. *Frontiers of Earth Science,* 7(7): 35-45.

Anthony, E.J. and Dolique, F. (2004). The influence of Amazon-derived mud banks on the morphology of sandy headland-bound beaches in Cayenne, French Guiana: a short- to long-term perspective. *Marine Geology,* 208(2): 249–264.

Anthony, E.J., Gardel, A. and Gratiot, N. (2014). Fluvial sediment supply, mud banks, cheniers and the morphodynamics of the coast of South America between the Amazon and Orinoco river mouths. *Geological Society, London, Special Publications* 388(1), 533–560.

Anthony, E.J., Gardel, A., Gratiot, N., Proisy, C., Allison, M.A., Dolique, F. and Fromard, F. (2010). The Amazon-influenced muddy coast of South America: A review of mud-bank–shoreline interactions. *Earth-Science Reviews*, 103(3): 99–121.

Anthony, E.J. and Gratiot, N. (2012). Coastal engineering and large-scale mangrove destruction in Guyana, South America: Averting an environmental catastrophe in the making. *Ecological Engineering*, 47(1): 268–273.

Arbelaez, P., Fowlkes, C. and Martin, M. (2007). *The Berkeley Segmentation Dataset and Benchmark.*Accessed online [13/01/2020]. Available at: https://www2.eecs.berkeley.edu/.

Bach, S., Binder, A., Montavon, G., Klauschen, F., Müller, K.-R. and Samek, W. (2015). On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation. *PLOS ONE*, 10(7): 1-46.

Ballère, M., Bouvet, A., Mermoz, S., Le Toan, T., Koleck, T., Bedeau, C., André, M., Forestier, E., Frison, P.L. and Lardeux, C. (2021). SAR data for tropical forest disturbance alerts in French Guiana: Benefit over optical imagery. *Remote Sensing of Environment*, 252(1): 1-15.

Barbier, E.B., Hacker, S.D., Kennedy, C., Koch, E.W., Stier, A.C. and Silliman, B.R. (2011). The value of estuarine and coastal ecosystem services. *Ecological Monographs*, 81(2): 169–193.

Barnes, E.A., Toms, B., Hurrell, J.W., Ebert-Uphoff, I., Anderson, C. and Anderson, D. (2020). Indicator Patterns of Forced Change Learned by an Artificial Neural Network. *Journal of Advances in Modeling Earth Systems*, 12(9): 1-18.

Barrett-Mold, C. and Burningham, H. (2010). Contrasting ecology of prograding coastal dunes on the northwest coast of Ireland. *Journal of Coastal Conservation*, 14(2): 81-90.

Bayram, B., Erdem, F., Akpinar, B., Ince, A.K., Bozkurt, S., Reis, H.C. and Seker, D.Z. (2017). The Efficiency of Random Forest Method for Shoreline Extraction from LANDSAT-8 and GOKTURK-2 Imageries. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 4(1): 141-145.

Belluco, E., Camuffo, M., Ferrari, S., Modenese, L., Silvestri, S., Marani, A. and Marani, M. (2006). Mapping salt-marsh vegetation by multispectral and hyperspectral remote sensing. *Remote Sensing of Environment*, 105(1): 54–67.

Berhane, T.M., Lane, C.R., Wu, Q., Autrey, B.C., Anenkhonov, O.A., Chepinoga, V.V. and Liu, H. (2018). Decision-Tree, Rule-Based, and Random Forest Classification of High-

Resolution Multispectral Imagery for Wetland Mapping and Inventory. *Remote Sensing*, 10(4): 580-606.

Beuzen, T., Splinter, K.D., Marshall, L.A., Turner, I.L., Harley, M.D. and Palmsten, M.L. (2018). Bayesian Networks in coastal engineering: Distinguishing descriptive and predictive applications. *Coastal Engineering*, 135(1): 16–30.

Bhargava, R., Sarkar, D. and Friess, D.A. (2021). A cloud computing-based approach to mapping mangrove erosion and progradation: Case studies from the Sundarbans and French Guiana. *Estuarine, Coastal and Shelf Science*, 248(1): 1-13.

Bhattacharya, J.P., Copeland, P., Lawton, T.F. and Holbrook, J. (2016). Estimation of source area, river paleo-discharge, paleoslope, and sediment budgets of linked deep-time depositional systems and implications for hydrocarbon potential. *Earth-Science Reviews*,153(1): 77–110.

Bishop-Taylor, R., Sagar, S., Lymburner, L., Alam, I. and Sixsmith, J. (2019). Sub-Pixel Waterline Extraction: Characterising Accuracy and Sensitivity to Indices and Spectra. *Remote Sensing*, 11(24): 2984-3001.

Boak, E.H. and Turner, I.L. (2005). Shoreline Definition and Detection: A Review. *Journal of Coastal Research*, 21(4): 688–703.

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1): 5-32.

Brock, J.C. and Purkis, S.J. (2009). The emerging role of LiDAR remote sensing in coastal research and resource management. *Journal of Coastal Research,* 53(10053): 1-5.

Brooks, H., Möller, I., Carr, S., Chirol, C., Christie, E., Evans, B., Spencer, K.L., Spencer, T. and Royse, K. (2021). Resistance of salt marsh substrates to near-instantaneous hydrodynamic forcing. *Earth Surface Processes and Landforms*, 46(1): 67-88.

Brooks, S.M. and Spencer, T. (2010). Temporal and spatial variations in recession rates and sediment release from soft rock cliffs, Suffolk coast, UK. *Geomorphology* 124(1-2): 26-41.

Brooks, S.M. and Spencer, T. (2012). Shoreline retreat and sediment release in response to accelerating sea level rise: Measuring and modelling cliffline dynamics on the Suffolk Coast, UK. *Global and Planetary Change*, 80(1): 165–179.

Brooks, S. and Spencer, T. (2019). Long-term trends, short-term shocks and cliff responses for areas of critical coastal infrastructure. In *Coastal Sediments 2019: Proceedings of the 9th International Conference*: 1179-1187.

Brooks, S.M. and Spencer, T., Christie, E.K. (2017). Storm impacts and shoreline recovery: Mechanisms and controls in the southern North Sea. *Geomorphology*, 283(1): 48–60.

Brownlee, J. (2019). *How do convolutional layers work in deep learning neural networks?* Accessed online [09/09/2021]. Available at: https://machinelearningmastery.com/convolutional-layers-for-deep-learning-neural-networks/

Brunier, G., Anthony, E.J., Gratiot, N. and Gardel, A. (2019). Exceptional rates and mechanisms of muddy shoreline retreat following mangrove removal. *Earth Surface Processes and Landforms,* 44(8): 1559–1571.

Buchhorn, M., Lesiv, M., Tsendbazar, N.-E., Herold, M., Bertels, L. and Smets, B. (2020). Copernicus Global Land Cover Layers—Collection 2. *Remote Sensing*, 12(6): 1044-1060.

Bullard, J.E., Ackerley, D., Millett, J., Chandler, J.H. and Montreuil, A.L. (2019). Post-storm geomorphic recovery and resilience of a prograding coastal dune system. *Environmental Research Communications*, 1(1): 1-12.

Bunting, P., Rosenqvist, A., Lucas, R.M., Rebelo, L.-M., Hilarides, L., Thomas, N., Hardy, A., Itoh, T., Shimada, M. and Finlayson, C.M. (2018). The Global Mangrove Watch—A New 2010 Global Baseline of Mangrove Extent. *Remote Sensing*, 10(10): 1669-1688.

Burke, L. and Ding, D. (2016). *Valuation of coastal protection near Paramaribo, Suriname.* Accessed online [01/07/2021]. Available at: http://d2ouvy59p0dg6k.cloudfront.net/downloads/3__16_02_project_peri_urban_coastal_protection_options_paramaribo___pre_final_report_wri.pdf

Burningham, H. and French, J. (2013). Is the NAO winter index a reliable proxy for wind climate and storminess in northwest Europe?. *International Journal of Climatology*, 33(8): 2036-2049.

Burningham, H. and French, J. (2017). Understanding coastal change using shoreline trend analysis supported by cluster-based segmentation. *Geomorphology,* 282(1): 131-149.

Burvingt, O.J.P., Masselink, G., Russell, P. and Scott, T. 2016. Beach response to consecutive extreme storms using LiDAR along the SW coast of England. *In:* Vila-Concejo, A., Bruce, E., Kennedy, D.M. and McCarroll, R.J. (eds.), *Proceedings of the 14th International Coastal Symposium* (Sydney, Australia). *Journal of Coastal Research*, 75(1): 1052 - 1056.

Buscombe, D. and Ritchie, A.C. (2018). Landscape Classification with Deep Neural Networks.*Geosciences*, 8(2): 244.

Caut, S., Guirlet, E. and Girondot, M. (2010). Effect of tidal overwash on the embryonic development of leatherback turtles in French Guiana. *Marine Environmental Research*, 69(4): 254-261.

Cefas. (2020). *WaveNet interactive map.* Accessed online [21/05/2020] Available at: http://wavenet.cefas.co.uk/Map.

Chai, T. and Draxler, R.R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature. *Geoscientific model development*, 7(3): 1247-1250.

Chen, B., Xiao, X., Li, X., Pan, L., Doughty, R., Ma, J., Dong, J., Qin, Y., Zhao, B., Wu, Z., Sun, R., Lan, G., Xie, G., Clinton, N. and Giri, C. (2017). A mangrove forest map of China in 2015: Analysis of time series Landsat 7/8 and Sentinel-1A imagery in Google Earth Engine cloud computing platform. *ISPRS Journal of Photogrammetry and Remote Sensing*,131: 104–120.

Chen, T. (2021). *Interpretability in Convolutional Neural Networks for Building Damage Classification in Satellite Imagery*. Accessed online [05/05/2021]. Available at: https://www.preprints.org/manuscript/202101.0053/v1

Choung, Y.J. and Jo, M.H. (2017). Comparison between a machine-learning-based method and a water-index-based method for shoreline mapping using a high-resolution satellite image acquired in Hwado Island, South Korea. *Journal of Sensors*, 2017(1):1-13.

Church, J.A., White, N.J., Coleman, R., Lambeck, K. and Mitrovica, J.X. (2004). Estimates of the Regional Distribution of Sea Level Rise over the 1950–2000 Period. *Journal of Climat*e, 17(13): 2609–2625.

Colenbrander, S., Lazar, L., Haddaoui, C., Godfrey, N., Lobo, A., Clarkson, H., Huxley, R., Parnell, S., Smith, B., Smith, S. and Altenburg, T. (2019). *Climate Emergency, Urban Opportunity: The unique and crucial roles of national governments.* Accessed online [14/02/2021]. Available at: https://urbantransitions.global/en/publication/climate-emergency-urban-opportunity/.

Collard, F., Ardhuin, F. and Chapron, B.(2009). Monitoring and analysis of ocean swell fields from space: New methods for routine observations. *Journal of Geophysical Research: Oceans*, 114(7): 1-10.

Committee on Climate Change. (2018). *Managing the coast in a changing climate.* Accessed online [14/03/2020]. Available at: https://www.theccc.org.uk/publication/managing-the-coast-in-a-changing-climate.

Cowell, P. J. and Thom, B. G. (1994). Morphodynamics of coastal evolution, in: Carter, R. W. G. & Woodroffe, C. D. (Eds), *Coastal Evolution, Late Quaternary shoreline morphodynamics.* Cambridge: Cambridge University Press, pp. 33 – 86.

Dai, C., Howat, I.M., Larour, E. and Husby, E. (2019). Coastline extraction from repeat high resolution satellite imagery. *Remote Sensing of Environment*, 229(1): 260–270.

DaSilva, M., Miot da Silva, G., Hesp, P.A., Bruce, D., Keane, R. and Moore, C. (2021). Assessing Shoreline Change using Historical Aerial and RapidEye Satellite Imagery (Cape Jaffa, South Australia). *Journal of Coastal Research,* 37(2): 468–483.

de Andrés, M., Barragán, J.M. and Scherer, M. (2018). Urban centres and coastal zone definition: Which area should we manage? *Land use policy,* 71(1): 121-128.

de Jong, S.M., Shen, Y., de Vries, J., Bijnaar, G., van Maanen, B., Augustinus, P. and Verweij, P. (2021). Mapping mangrove dynamics and colonization patterns at the Suriname coast using historic satellite data and the LandTrendr algorithm. *International Journal of Applied Earth Observation and Geoinformation*, 97(1): 1-11.

de Vries, J., van Maanen, B., Ruessink, G., Verweij, P.A. and de Jong, S.M. (2021). Unmixing water and mud: Characterizing diffuse boundaries of subtidal mud banks from individual satellite observations. *International Journal of Applied Earth Observation and Geoinformation*, 95(1): 1-20.

Defra, (2006). *Shoreline management plan guidance- Volume 1: Aims and requirements-March 2001.* Accessed online [16/06/2021]. Available at: www.defra.gov.uk

Demir, N., Oy, S., Erdem, F., Şeker, D.Z. and Bayram, B. (2017). Integrated shoreline extraction approach with use of Rasat MS and SENTINEL-1A SAR Images. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 4(1): 445-449.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K. and Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database, in: 2009 IEEE Conference on Computer Vision and Pattern Recognition. Presented at the 2009 *IEEE Conference on Computer Vision and Pattern Recognition*: 248–255.

Dominici, D., Zollini, S., Alicandro, M., Della Torre, F., Buscema, P.M. and Baiocchi, V. (2019). High resolution satellite images for instantaneous shoreline extraction using new enhancement algorithms. *Geosciences*, 9(3): 123.

Dueben, P.D. and Bauer, P. (2018). Challenges and design choices for global weather and climate models based on machine learning. *Geoscientific Model Development*, 11(10): 3999–4009.

Dytham, C. (2011). *Choosing and using statistics: a biologist's guide*. John Wiley & Sons: Oxford, U.K.

Ebert-Uphoff, I., Hilburn, K. (2020). Evaluation, Tuning, and Interpretation of Neural Networks for Working with Images in Meteorological Applications. *Bulletin of the American Meteorological Society*, 101(12): E2149–E2170.

Elgohary, T., Mubasher, A. and Salah, H. (2017). Significant deep wave height prediction by using support vector machine approach (Alexandria as case of study). *International Journal of Current Engineering and Technology*, 7(1): 135-143.

Elkafrawy, S.B., Basheer, M.A., Mohamed, H.M. and Naguib, D.M. (2021). Applications of remote sensing and GIS techniques to evaluate the effectiveness of coastal structures along Burullus headland-Eastern Nile Delta, Egypt. *The Egyptian Journal of Remote Sensing and Space Science*, 24(2): 247–254.

Elnabwy, M.T., Elbeltagi, E., El Banna, M.M., Elshikh, M.M., Motawa, I. and Kaloop, M.R. (2020). An Approach Based on Landsat Images for Shoreline Monitoring to Support Integrated Coastal Management—A Case Study, Ezbet Elborg, Nile Delta, Egypt. *International Society for Photogrammetry and Remote Sensing International Journal of Geo-Information* 9(4): 199-218.

Environment Agency, (2010). *Coastal Morphology Report.*Accessed online [09/06/2021]. Available at: https://coastalmonitoring.org/anglian/analysis_programme/Coastal%20Morphology%20Report%20Southwold%20to%20Benacre%20Denes%20Suffolk%20RP016S2010.pdf

Environment Agency, (2016). *The costs and impacts of the winter 2013 to 2014 floods.* Accessed online [10/12/2019]. Available at: https://assets.publishing.service.gov.uk/media/60354990e90e0740b7caac90/The_costs_and_impacts_of_the_winter_2013_to_2014_floods_-_non_technical_report.pdf

Environment Agency, (2020a). *Vertical Aerial Photography.* Accessed online [19/06/2020]. Available at:https://data.gov.uk/dataset.

Environment Agency, (2020b). *National Flood and Coastal Erosion Risk Management Strategy for England.* Accessed online [01/08/2021]. Available at: https://www.gov.uk/government/publications/national-flood-and-coastal-erosion-risk-management-strategy-for-england--2

Environment Agency, (2021). *Asset Information and Maintenance Programme (AIMs)*. Accessed online [18/08/2021]. Available at: https://environment.data.gov.uk/asset-management/index.html

Environment Agency, (n.d). *East Coast Coastal Group: Coastal Monitoring.* Accessed online [13/07/2020]. Available at: http://www.eacg.org.uk/default_monitoring.asp

European Space Agency, (2021). *SkySat Mission: SkySat Imagery and SkySat Data.* Accessed online [09/08/2021]. Available at: https://earth.esa.int/eogateway/missions/skysat

FAO, (2007). The World's Mangroves 1980-2005. *FAO Forestry Paper*, 153: 77.

Farris, A.S., Defne, Z. and Ganju, N.K. (2019). Identifying Salt Marsh Shorelines from Remotely Sensed Elevation Data and Imagery.*Remote Sensing*, 11(15): 1795-1012.

Feilhauer, H., Zlinszky, A., Kania, A., Foody, G.M., Doktor, D., Lausch, A. and Schmidtlein, S. (2020). Let your maps be fuzzy!—Class probabilities and floristic gradients as alternatives to crisp mapping for remote sensing of vegetation. *Remote Sensing in Ecology and Conservation,* 7(2): 292-305.

Ferreira, O., Garcia, T., Matias, A., Taborda, R. and Dias, J.A. (2006). An integrated method for the determination of set-back lines for coastal erosion hazards on sandy shores. *Continental Shelf Research, 26*(9): 1030-1044.

Floodlist. (2017). *UK – Storm Surge Causes Flooding Along England's East Coast.* Accessed online [17/06/2020]. Available at: www.http://floodlist.com/europe/united-kingdom.

Foody, G.M., Muslim, A.M. and Atkinson, P.M. (2005). Super-resolution mapping of the waterline from remotely sensed data.*International Journal of Remote Sensing,* 26(24): 5381-5392.

Frantz, D., Haß, E., Uhl, A., Stoffels, J. and Hill, J. (2018). Improvement of the Fmask algorithm for Sentinel-2 images: Separating clouds from bright surfaces based on parallax effects. *Remote Sensing of Environment,* 215(1): 471–481.

Froidefond, J.M., Pujos, M. and Andre, X. (1988). Migration of mud banks and changing coastline in French Guiana. *Marine Geology,* 84(1-2): 19–30.

Fromard, F., Puig, H., Mougin, E., Marty, G., Betoulle, J.L. and Cadamuro, L. (1998). Structure, above-ground biomass and dynamics of mangrove ecosystems: new data from French Guiana. *Oecologia,* 115(1): 39–53.

Fromard, F., Vega, C. and Proisy, C. (2004). 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. *Marine Geology*208: 265–280.

Fu, W., Ma, J., Chen, P. and Chen, F. (2020). Remote Sensing Satellites for Digital Earth, in: Guo, H., Goodchild, M.F., Annoni, A. (Eds.), *Manual of Digital Earth*. Springer, Singapore, pp. 55–123.

Gandhi, G.M., Parthiban, S., Thummalu, N. and Christy, A. (2015). Ndvi: Vegetation change detection using remote sensing and gis–A case study of Vellore District. *Procedia Computer Science,* 57(1): 1199-1210.

Gandhi, G.M., Parthiban, S., Thummalu, N. and Christy, A. (2015). Ndvi: Vegetation Change Detection Using Remote Sensing and Gis – A Case Study of Vellore District.

Procedia Computer Science, 3rd International Conference on Recent Trends in Computing 2015 (ICRTC-2015),57(1): 1199–1210.

García-Rubio, G., Huntley, D. and Russell, P. (2015). Evaluating shoreline identification using optical satellite images. *Marine Geology*, 359(1): 96-105.

Gardel, A. and Gratiot, N. (2005). A Satellite Image–Based Method for Estimating Rates of Mud Bank Migration, French Guiana, South America. *Journal of Coastal Research,* 21(4): 720–728.

Gardel, A. and Gratiot, N. (2006). Monitoring of Coastal Dynamics in French Guiana from 16 Years of SPOT Satellite Images. *Journal of Coastal Research*, 39(1): 1502–1505.

Gardel, A., Proisy, C., Lesourd, S., Philippe, S., Caillaud, J., Gontharet, S., Anthony, E.J. and Brutier, L. (2009). A Better Understanding of Mud Cracking Processes Gained From in Situ Measurements on an Intertidal Mudflat in French Guiana. *Journal of Coastal Research*, 56(1): 424–428.

General Bureau of Statistics Suriname, (2021). *The Census.* Accessed online [29/06/2021]. Available at: https://statistics-suriname.org/en/the-census/.

Genovese, G., Vignolles, C., Nègre, T. and Passera, G. (2001). A methodology for a combined use of normalised difference vegetation index and CORINE land cover data for crop yield monitoring and forecasting. A case study on Spain. *Agronomie,* 21(1): 91-111.

Gentine, P., Pritchard, M., Rasp, S., Reinaudi, G. and Yacalis, G. (2018). Could Machine Learning Break the Convection Parameterization Deadlock? *Geophysical Research Letters*, 45(11): 5742–5751.

Ghorbani, A., Abid, A. and Zou, J. (2018). Interpretation of Neural Networks is Fragile. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(1): 3681-3688.

Giri, C., Ochieng, E., Tieszen, L.L., Zhu, Z., Singh, A., Loveland, T., Masek, J. and Duke, N. (2011). Status and distribution of mangrove forests of the world using earth observation satellite data. *Global Ecology and Biogeography*, 20(1): 154–159.

Goldstein, E.B., Coco, G. and Plant, N.G. (2019). A review of machine learning applications to coastal sediment transport and morphodynamics. *Earth-Science Reviews,* 194(1): 97–108.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D. and Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202(1): 18–27.

Government of Guyana, (2021). *Census: What is the census?.* Accessed online [29/06/2021]. Available at:https://statisticsguyana.gov.gy/census/.

Gratiot, N., Anthony, E.J., Gardel, A., Gaucherel, C., Proisy, C. and Wells, J.T. (2008). Significant contribution of the 18.6 year tidal cycle to regional coastal changes. *Nature Geoscience*, 1(3): 169–172.

Grzegorzewski, A.S., Cialone, M.A. and Wamsley, T.V. (2011). Interaction of barrier islands and storms: implications for flood risk reduction in Louisiana and Mississippi. *Journal of Coastal Research*, 1(59): 156-164.

Gu, T., Dolan-Gavitt, B. and Garg, S. (2019). BadNets: Identifying Vulnerabilities in the Machine Learning Model Supply Chain. *Cryptography and Security*¸1(1): 1708-1721.

Gupta, U., Joshi, P., Bodani, P., Oza, S., Rajak, D. and Oza, M. (2019). Development of a Web-GIS based system for safer ship navigation in Antarctic region using open source technologies. *Journal of Geomatics*, 13(2): 203-208.

Hagenaars, G., de Vries, S., Luijendijk, A.P., de Boer, W.P. and Reniers, A.J. (2018). On the accuracy of automated shoreline detection derived from satellite imagery: A case study of the sand motor mega-scale nourishment. *Coastal Engineering*, 133(1): 113-125.

Hamdi, Z.M., Brandmeier, M. and Straub, C. (2019). Forest Damage Assessment Using Deep Learning on High Resolution Remote Sensing Data. *Remote Sensing*, 11(17): 1976.

Hapke, C.J., Himmelstoss, E.A., Kratzmann, M.G., List, J.H. and Thieler, E.R. (2010). *National assessment of shoreline change: Historical shoreline change along the New England and Mid-Atlantic coasts. US Geological Survey.*Accessed online [03/09/2019]. Available at: https://tamug-ir.tdl.org/bitstream/handle/1969.3/29280/ofr2010-1118_report_508_rev042312.pdf?sequence=1

Hardy, J.R. (1964). The movement of beach material and wave action near Blakeney Point, Norfolk. *Transactions and Papers (Institute of British Geographers)*, 1(34): 53-69.

Hasan, H., Shafri, H.Z.M. and Habshi, M. (2019). A Comparison Between Support Vector Machine (SVM) and Convolutional Neural Network (CNN) Models for Hyperspectral Image Classification. *IOP Conference Series: Earth and Environmental Science,* 357(1): 012035-012046.

Hein, C.J., Fallon, A.R., Rosen, P., Hoagland, P., Georgiou, I.Y., FitzGerald, D.M., Morris, M., Baker, S., Marino, G.B. and Fitzsimons, G. (2019). Shoreline Dynamics Along a Developed River Mouth Barrier Island: Multi-Decadal Cycles of Erosion and Event-Driven Mitigation. *Frontiers in Earth Science*, 7(1): 103

Hickey, C. and Weis, T. (2012). The challenge of climate change adaptation in Guyana. *Climate and Development*, 4(1): 66–74.

Hilburn, K.A., Ebert-Uphoff, I. and Miller, S.D. (2020). Development and Interpretation of a Neural-Network-Based Synthetic Radar Reflectivity Estimator Using GOES-R Satellite Observations. *Journal of Applied Meteorology and Climatology,* 60(1): 3–21.

Hinkel, J., Lincke, D., Vafeidis, A.T., Perrette, M., Nicholls, R.J., Tol, R.S., Marzeion, B., Fettweis, X., Ionescu, C. and Levermann, A. (2014). Coastal flood damage and adaptation costs under 21st century sea-level rise. *Proceedings of the National Academy of Sciences*, 111(9): 3292-3297.

Hoeser, T., Bachofer, F. and Kuenzer, C. (2020). Object Detection and Image Segmentation with Deep Learning on Earth Observation Data: A Review—Part II: Applications. *Remote Sensing*, 12(18): 3053-3100.

Houborg, R. and McCabe, M.F. (2018). A Cubesat enabled Spatio-Temporal Enhancement Method (CESTEM) utilizing Planet, Landsat and MODIS data. *Remote Sensing of Environment*, 209(1): 211–226.

Hu, T., Zhang, Y., Su, Y., Zheng, Y., Lin, G. and Guo, Q. (2020). Mapping the Global Mangrove Forest Aboveground Biomass Using Multisource Remote Sensing Data. *Remote Sensing*, 12(10): 1690.

Ii, D.J.G., Haupt, S.E., Nychka, D.W. and Thompson, G. (2019). Interpretable Deep Learning for Spatial Analysis of Severe Hailstorms. *Monthly Weather Review,* 147(8): 2827–2845.

Jäger, W.S., Christie, E.K., Hanea, A.M., den Heijer, C. and Spencer, T. (2018). A Bayesian network approach for coastal risk analysis and decision making. *Coastal Engineering*, 134(1): 48–61.

Jana, A., Maiti, S. and Biswas, A. (2015). Seasonal change monitoring and mapping of coastal vegetation types along Midnapur-Balasore Coast, Bay of Bengal using multi-temporal landsat data. *Modeling Earth Systems and Environment,* 2(1): 1-12.

Jergensen, G.E., McGovern, A., Lagerquist, R. and Smith, T. (2020). Classifying Convective Storms Using Machine Learning. *Weather and Forecasting,* 35(2): 537–559.

Jolivet, M., Gardel, A. and Anthony, E.J. (2019). Multi-decadal Changes on the Mud-dominated Coast of Western French Guiana: Implications for Mesoscale Shoreline Mobility, River-mouth Deflection, and Sediment Sorting. *Journal of Coastal Research,* 88(SI): 185–194.

Kaieteur News, (2017). *Authorities order Shell Beach closed over floods, erosion.* Accessed online [17/05/2021]. Available at: https://www.kaieteurnewsonline.com/2017/04/23/authorities-order-shell-beach-closed-over-floods-erosion/.

Karpatne, A., Jiang, Z., Vatsavai, R.R., Shekhar, S. and Kumar, V. (2016). Monitoring Land-Cover Changes: A Machine-Learning Perspective. *IEEE Geoscience and Remote Sensing Magazine*, 4(2): 8–21.

Kashinath, K., Mustafa, M., Albert, A., Wu, J.-L., Jiang, C., Esmaeilzadeh, S., Azizzadenesheli, K., Wang, R., Chattopadhyay, A., Singh, A., Manepalli, A., Chirila, D., Yu, R., Walters, R., White, B., Xiao, H., Tchelepi, H.A., Marcus, P., Anandkumar, A. and Hassanzadeh, P. (2021). Physics-informed machine learning: case studies for weather and climate modelling. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379(2194): 20200093-20200129.

Katiyar, S.K. and Arun, P.V. (2014). Comparative analysis of common edge detection techniques in context of object extraction.*Computer Vision and Pattern Recognition*,1405(1): 1-12.

Ketkar, N. (2017). Introduction to Keras, in: Ketkar, N. (Ed.), *Deep Learning with Python: A Hands-on Introduction*. Apress, Berkeley, CA, pp. 97–111.

Kim, J., Han, H., Johnson, L.E., Lim, S. and Cifelli, R. (2019). Hybrid machine learning framework for hydrological assessment. *Journal of Hydrology*, 577(1): 123913-123924.

Klemas, V. (2013). Remote sensing of emergent and submerged wetlands: an overview. *International Journal of Remote Sensing*, 34(18): 6286–6320.

Kokkinos, I. (2015). Pushing the boundaries of boundary detection using deep learning. *Computer Vision and Pattern Recognition*,1511(1): 1-12.

Komar, P.D. (1998) *Beach processes and sedimentation* (Second edition). New Jersey: Prentice Hall, 544pp.

Kron, W. (2013). Coasts: the high-risk areas of the world. *Natural Hazards*, 66(3): 1363–1382.

Kuleli, T., Guneroglu, A., Karsli, F. and Dihkan, M. (2011). Automatic detection of shoreline change on coastal Ramsar wetlands of Turkey. *Ocean Engineering*, 38(10): 1141–1149.

Kulp, S. and Strauss, B.H. (2017). Rapid escalation of coastal flood exposure in US municipalities from sea level rise. *Climatic Change*, 142(3-4): 477-489.

Leatherman, S.P. (2003). Shoreline change mapping and management along the US East Coast. *Journal of Coastal Research*, 38(1): 5-13.

Lee, Y., Kummerow, C.D. and Ebert-Uphoff, I. (2020a). Applying machine learning methods to detect convection usingGOES-16 ABI data (preprint). Clouds/Remote Sensing/Data Processing and Information Retrieval.

Lee, Y., Kummerow, C.D. and Zupanski, M. (2020b). A simplified method for the detection of convection using high resolution imagery from GOES-16. *Atmospheric Measurement Techniques Discussions,* 14(5): 1–26.

Lefebvre, J.P., Dolique, F. and Gratiot, N. (2004). Geomorphic evolution of a coastal mudflat under oceanic influences: an example from the dynamic shoreline of French Guiana. *Marine Geology*, 208: 191–205.

Lewis, R.R. (2005). Ecological engineering for successful management and restoration of mangrove forests. *Ecological Engineering,* 24(4): 403–418.

Li, R., Di, K. and Ma, R. (2001). A comparative study of shoreline mapping techniques. *GIS for coastal zone management,* pp.53-60.

Li, W. and Gong, P. (2016). Continuous monitoring of coastline dynamics in western Florida with a 30-year time series of Landsat imagery. *Remote Sensing of Environment,* 179(1): 196-209.

Lincke, D. and Hinkel, J. (2018). Economically robust protection against 21st century sea-level rise. *Global Environmental Change*, 51(1): 67-73.

Liu, C (2018). HED: Keras implementation. Github repository https://github.com/lc82111.

Liu, H. and Jezek, K.C. (2004). Automated extraction of coastline from satellite imagery by integrating Canny edge detection and locally adaptive thresholding methods. *International Journal of Remote Sensing*, 25(5): 937–958.

Liu, H., Luo, J., Sun, Y., Xia, L., Wu, W., Yang, H., Hu, X. and Gao, L. (2019). Contour-oriented Cropland Extraction from High Resolution Remote Sensing Imagery Using Richer Convolution Features Network. *In 2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics),*8(1): 1-6.

Liu, M., Yu, T., Gu, X., Sun, Z., Yang, J., Zhang, Z., Mi, X., Cao, W. and Li, J. (2020). The Impact of Spatial Resolution on the Classification of Vegetation Types in Highly Fragmented Planting Areas Based on Unmanned Aerial Vehicle Hyperspectral Images. *Remote Sensing,* 12(1): 146.

Liu, X.Y., Jia, R.S., Liu, Q.M., Zhao, C.Y. and Sun, H.M. (2019). Coastline Extraction Method Based on Convolutional Neural Networks—A Case Study of Jiaozhou Bay in Qingdao, China. *IEEE Access,* 7(1): 180281-180291.

Lloyd, C.T., Chamberlain, H., Kerr, D., Yetman, G., Pistolesi, L., Stevens, F.R., Gaughan, A.E., Nieves, J.J., Hornby, G., MacManus, K., Sinha, P., Bondarenko, M., Sorichetta, A. and Tatem, A.J. (2019). Global spatio-temporally harmonised datasets for producing high-resolution gridded population distribution datasets. *Big Earth Data,* 3(2): 108–139

Loveland, T.R. and Irons, J.R. (2016). Landsat 8: The plans, the reality, and the legacy. *Remote Sensing of Environment*, 185(1): 1-6.

Lu, T., Ming, D., Lin, X., Hong, Z., Bai, X. and Fang, J. (2018). Detecting building edges from high spatial resolution remote sensing imagery using richer convolution features network. *Remote Sensing*,10(9): 1496- 1515.

Luijendijk, A., Hagenaars, G., Ranasinghe, R., Baart, F., Donchyts, G., Aarninkhof, S. (2018). The State of the World's Beaches. *Scientific Reports*, 8(1): 1-11.

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G. and Johnson, B.A. (2019). Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152(1): 166–177.

Marmanis, D., Schindler, K., Wegner, J.D., Galliani, S., Datcu, M. and Stilla, U. (2018). Classification with an edge: Improving semantic image segmentation with boundary detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, 135(1): 158-172.

Marta, S. (2018). *Planet imagery product specifications.* Accessed online [05/02/2019]. Available at: https://www.planet.com/products/satellite-imagery/files/1610.06_Spec%20Sheet_Combined_Imagery_Product_Letter_ENGv1.pdf

Masek, J.G., Wulder, M.A., Markham, B., McCorkel, J., Crawford, C.J., Storey, J. and Jenstrom, D.T. (2020). Landsat 9: Empowering open science and applications through continuity. *Remote Sensing of Environment*, 248(1): 1-13.

Maskrey, S.A., Priest, S. and Mount, N.J. (2019). Towards evaluation criteria in participatory flood risk management. *Journal of Flood Risk Management*, 12(2): 1-14.

Masselink, G. and van Heteren, S. (2014). Response of wave-dominated and mixed-energy barriers to storms. *Marine Geology*, 352(1): 321–347.

McCarthy, M.J., Jessen, B., Barry, M.J., Figueroa, M., McIntosh, J., Murray, T., Schmid, J. and Muller-Karger, F.E. (2020). Mapping hurricane damage: A comparative analysis of satellite monitoring methods. *International Journal of Applied Earth Observation and Geoinformation*, 91(1): 1-9.

McFeeters, S.K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17(7): 1425-1432.

McGovern, A., Lagerquist, R., Gagne, D.J., Jergensen, G.E., Elmore, K.L., Homeyer, C.R. and Smith, T. (2019). Making the Black Box More Transparent: Understanding the Physical Implications of Machine Learning. *Bulletin of the American Meteorological Society*, 100(11): 2175–2199.

McGranahan, G., Balk, D. and Anderson, B. (2007). The rising tide: assessing the risks of climate change and human settlements in low elevation coastal zones. *Environment and Urbanization*, 19(1): 17–37.

McLean, R.F. and Kirk, R.M. (1969) Relationship between grain size, size-sorting and foreshore slope on mixed sand-shingle beaches. *New Zealand Journal of Geology and Geophysics* 12(1), 138-155.

McLoughlin, S.M., Wiberg, P.L., Safak, I. and McGlathery, K.J. (2015). Rates and Forcing of Marsh Edge Erosion in a Shallow Coastal Bay. *Estuaries and Coasts*, 38(2): 620–638.

Medina Machín, A., Marcello, J., Hernández-Cordero, A.I., Martín Abasolo, J., Eugenio, F., 2019. Vegetation species mapping in a coastal-dune ecosystem using high resolution satellite imagery. *GIScience & Remote Sensing*, 56(2): 210–232.

Mentaschi, L., Vousdoukas, M.I., Pekel, J.-F., Voukouvalas, E. and Feyen, L. (2018). Global long-term observations of coastal erosion and accretion. *Scientific Reports*, 8(1): 12876.

Miller, T.E., Gornish, E.S. and Buckley, H.L. (2010). Climate and coastal dune vegetation: disturbance, recovery, and succession. *Plant Ecology*, 206(1): 97-104.

Mohajerani, Y., Wood, M., Velicogna, I. and Rignot, E. (2019). Detection of glacier calving margins with convolutional neural networks: A case study. *Remote Sensing* 11(1): 74-87.

Möller, I. (2006). Quantifying saltmarsh vegetation and its effect on wave height dissipation: Results from a UK East coast saltmarsh. *Estuarine, Coastal and Shelf Science*, 69(3): 337–351.

Möller, I. and Spencer, T. (2002). Wave dissipation over macro-tidal saltmarshes: Effects of marsh edge typology and vegetation change. *Journal of Coastal Research*, 36(10036): 506–521.

Montavon, G., Samek, W. and Müller, K.-R. (2018). Methods for interpreting and understanding deep neural networks. *Digital Signal Processing*, 73(1): 1–15.

Moore, L.J., Ruggiero, P. and List, J.H. (2006). Comparing mean high water and high water line shorelines: should proxy-datum offsets be incorporated into shoreline change analysis? *Journal of Coastal Research*, 224(4): 894-905.

Mountrakis, G., Im, J. and Ogole, C. (2011). Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3): 247–259.

Nagle, N.N., Buttenfield, B.P., Leyk, S. and Spielman, S. (2014). Dasymetric Modeling and Uncertainty. *Annals of the Association of American Geographers*, 104(1): 80–95.

Nassar, K., Mahmod, W.E., Fath, H., Masria, A., Nadaoka, K. and Negm, A. (2019). Shoreline change detection using DSAS technique: Case of North Sinai coast, Egypt. *Marine Georesources & Geotechnology,* 37(1): 81–95.

National Oceanic and Atmospheric Administration, (2021). *Environmental Monitoring Centre. NOAA WaveWatch III.* Accessed online [09/08/2021]. Available at: https://polar.ncep.noaa.gov/waves/

National Oceanography Centre, (2021a). *Permanent Service for Mean Sea Level: Ile Royale.* Accessed online [15/05/2021]. Available at: https://www.psmsl.org/data/obtaining/stations/2012.php

National Oceanography Centre, (2021b). *UK tide gauge network*. Accessed online [09/08/2021]. Available at: https://www.bodc.ac.uk/data/hosted_data_systems/sea_level/uk_tide_gauge_network/

Neumann, B., Vafeidis, A.T., Zimmermann, J. and Nicholls, R.J. (2015). Future Coastal Population Growth and Exposure to Sea-Level Rise and Coastal Flooding - A Global Assessment. *PLOS ONE*, 10(3): 1-34.

NOAA. (2021). *Cold and warm episodes by season.* Accessed online [15/5/2021]. Available at: https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php

Oppenheimer, M., B.C. Glavovic , J. Hinkel, R. van de Wal, A.K. Magnan, A. Abd-Elgawad, R. Cai, M. Cifuentes-Jara, R.M. DeConto, T. Ghosh, J. Hay, F. Isla, B. Marzeion, B. Meyssignac, and Z. Sebesvari, (2019). Sea Level Rise and Implications for Low-Lying Islands, Coasts and Communities. In: H.-O. Pörtner, D.C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegría, M. Nicolai, A. Okem, J. Petzold, B. Rama and N.M. Weyer (eds.). *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate.* Cambridge University Press, Cambridge, UK, pp. 321-445.

ORNL DAAC. (2018). *MODIS and VIIRS Land Products Global Subsetting and Visualization Tool.ORNL DAAC, Oak Ridge, Tennessee, USA.* Accessed Online [01/06/2021]. Available at: https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1379

Pardo-Pascual, J.E., Almonacid-Caballer, J., Ruiz, L.A. and Palomar-Vázquez, J. (2012). Automatic extraction of shorelines from Landsat TM and ETM+ multi-temporal images with subpixel precision. *Remote Sensing of Environment*, 123(1): 1-11.

Pardo-Pascual, J.E., Sánchez-García, E., Almonacid-Caballer, J., Palomar-Vázquez, J.M., Priego de los Santos, E., Fernández-Sarría, A. and Balaguer-Beser, Á. (2018). Assessing the Accuracy of Automatically Extracted Shorelines on Microtidal Beaches from Landsat 7, Landsat 8 and Sentinel-2 Imagery. *Remote Sensing,* 10(2): 326.

Parente, C., Alcaras, E., Errico, A., Falchi, U. and Vallario, A. (2019). Coastline Extraction from Optical Satellite Imagery and Accuracy Evaluation. *In International Workshop on R3 in Geomatics: Research, Results and Review*. Springer, Cham, pp. 336-349.

Park, S.-J. and Lee, D.-K. (2020). Prediction of coastal flooding risk under climate change impacts in South Korea using machine learning algorithms. *Environmental Research Letters*, 15(9): 1-11.

Payo, A., Jigena Antelo, B., Hurst, M., Palaseanu-Lovejoy, M., Williams, C., Jenkins, G., Lee, K., Favis-Mortlock, D., Barkwith, A. and Ellis, M.A. (2018). Development of an automatic delineation of cliff top and toe on very irregular planform coastlines (CliffMetrics v1.0). *Geoscientific Model Development*, 11(10): 4317–4337.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V. and Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. *Journal of machine Learning research*, 12(1): 2825-2830.

Pekel, J.-F., Cottam, A., Gorelick, N. and Belward, A.S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, 540(7633): 418–422.

Penning-Rowsell, E.C., de Vries, W.S., Parker, D.J., Zanuttigh, B., Simmonds, D., Trifonova, E., Hissel, F., Monbaliu, J., Lendzion, J., Ohle, N., Diaz, P. and Bouma, T. (2014). Innovation in coastal risk management: An exploratory analysis of risk governance issues at eight THESEUS study sites. *Coastal Engineering*, 87(1): 210–217.

Pielke, R. A. (2008). Normalised hurricane damage in the United States: 1900–2005. *Natural Hazards Review*, 9(1): 29–42.

Planet Team (2017). Planet Application Program Interface: In *Space for Life on Earth*. San Francisco, CA, pp. 1-10.

Plaziat, J.-C. and Augustinus, P.G.E.F. (2004). Evolution of progradation/erosion along the French Guiana mangrove coast: a comparison of mapped shorelines since the 18th century with Holocene data. *Marine Geology*, 208(2-4): 127–143.

Pollard, J. A., Brooks, S.M. and Spencer, T. (2019a). Harmonising topographic & remotely sensed datasets, a reference dataset for shoreline and beach change analysis. *Nature Scientific Data*, 6(1): 1-42.

Pollard, J.A., Spencer, T. and Brooks, S.M. (2019b). The interactive relationship between coastal erosion and flood risk. *Progress in Physical Geography: Earth and Environment,* 43(4): 574-585.

Pollard, J.A., Spencer, T., Brooks, S.M., Christie, E.K. and Möller, I. (2020). Understanding spatio-temporal barrier dynamics through the use of multiple shoreline proxies. *Geomorphology*, 354(1): 1-14.

Pollard, J.A., Spencer, T. and Jude, S. (2018). Big Data Approaches for coastal flood risk assessment and emergency response. *Wiley Interdisciplinary Reviews: Climate Change* 9(5), 1-14.

Pontee, N. (2013). Defining coastal squeeze: A discussion. *Ocean & Coastal Management*, 84(1): 204–207.

Proisy, C., Walcker, R., Blanchard, E., Gardel, A. and Anthony, E.J. (2021). Chapter 2 - Mangroves: a natural early-warning system of erosion on open muddy coasts in French Guiana, in: Sidik, F. and Friess, D.A. (Eds.), *Dynamic Sedimentary Environments of Mangrove Coasts*. Elsevier, pp. 47–66.

Pugh, D. and Woodworth, P. (2014). *Sea-level science: understanding tides, surges, tsunamis and mean sea-level changes*. Cambridge University Press, Cambridge.

Pye, K. and Blott, S.J. (2006). Coastal processes and morphological change in the Dunwich-Sizewell area, Suffolk, UK. *Journal of Coastal Research*, 22(3): 453-473.

Rahman, A.F., Dragoni, D. and El-Masri, B. (2011). Response of the Sundarbans coastline to sea level rise and decreased sediment flow: A remote sensing assessment. *Remote Sensing of Environment*, 115(12): 3121-3128.

Ramieri, E., Hartley, A., Barbanti, A., Santos, F.D., Gomes, A., Laihonen, P., Marinova, N. and Santini, M. (2011). Methods for assessing coastal vulnerability to climate change. *ETC CCA technical paper*, 1(2011):1-93.

Ranasinghe, R., A. C. Ruane, R. Vautard, N. Arnell, E. Coppola, F. A. Cruz, S. Dessai, A. S. Islam, M. Rahimi, D. Ruiz Carrascal, J. Sillmann, M. B. Sylla, C. Tebaldi, W. Wang, and R. Zaaboul, (2021). Climate Change Information for Regional Impact and for Risk Assessment, in: MassonDelmotte, V., P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu and B. Zhou (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press. In Press.

Rasp, S., Dueben, P.D., Scher, S., Weyn, J.A., Mouatadid, S. and Thuerey, N. (2020). WeatherBench: A Benchmark Data Set for Data-Driven Weather Forecasting. *Journal of Advances in Modeling Earth Systems*, 12(11): 1-17.

Ren, X. (2008). Multi-scale improves boundary detection in natural images. In European conference on computer vision. Springer, Berlin, Heidelberg, pp. 533-545.

Rogan, J., Franklin, J., Stow, D., Miller, J., Woodcock, C. and Roberts, D. (2008). Mapping land-cover modifications over large areas: A comparison of machine learning algorithms. *Remote Sensing of Environment*, 112(1): 2272–2283.

Rogers, M.S.J., Bithell, M., Brooks, S.M. and Spencer, T. (2021). VEdge_Detector: automated coastal vegetation edge detection using a convolutional neural network. *International Journal of Remote Sensing*, 42(3): 4805–4835.

Roy, S. (2017). Planet to Google Earth Engine Pipeline (Command Line Interface). Planet-GEE-Pipeline-CLI.

Schmidt, K.S., Skidmore, A.K., Kloosterman, E.H., van Oosten, H., Kumar, L. and Janssen, J.A.M. (2004). Mapping Coastal Vegetation Using an Expert System and Hyperspectral Imagery. *Photogrammetric Engineering & Remote Sensing,* 70(1): 703–715.

Scott, T., Masselink, G., O'Hare, T., Saulter, A., Poate, T., Russell, P., Davidson, M. and Conley, D. (2016). The extreme 2013/2014 winter storms: Beach recovery along the southwest coast of England. *Marine Geology,* 382(2): 224–241.

Scrimshaw, M.D., Bubb, J.M. and Lester, J.N. (1996). Organochlorine contamination of UK Essex coast salt marsh sediments. *Journal of Coastal Research*, 12(1): 246-255.

Shahinfar, S., Meek, P. and Falzon, G. (2020). How many images do I need? Understanding how sample size per class affects deep learning model performance metrics for balanced designs in autonomous wildlife monitoring. *Ecological Informatics,* 57(1): 1-14.

Sharma, C. (2017). *Convolutional Neural Networks in Python with Keras.* Accessed online [09/09/2021]. Available at: https://www.datacamp.com/community/tutorials/convolutional-neural-networks-python

Shin, M.C., Goldgof, D.B. and Bowyer, K.W. (2001). Comparison of edge detector performance through use in an object recognition task.*Computer Vision and Image Understanding*, 84(1): 160-178.

Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *Computer Vision and Pattern Recognition,* 1409(1): 1-14.

Spencer, T., Brooks, S.M., Evans, B.R., Tempest, J.A. and Möller, I. (2015). Southern North Sea storm surge event of 5 December 2013: water levels, waves and coastal impacts. *Earth-Science Reviews,* 146(1): 120-145.

Spencer, T., Möller, I. and Reef, R. (2016). Mangrove Systems and Environments, in: *Reference Module in Earth Systems and Environmental Sciences*. Elsevier, Netherlands, pp.1-34.

Splinter, K.D., Carley, J.T., Golshani, A. and Tomlinson, R. (2014). A relationship to describe the cumulative impact of storm clusters on beach erosion. *Coastal Engineering,* 83(1): 49–55.

Splinter, K.D., Harley, M.D. and Turner, I.L. (2018). Remote Sensing Is Changing Our View of the Coast: Insights from 40 Years of Monitoring at Narrabeen-Collaroy, Australia. *Remote Sensing,* 10(11): 1744-1769.

Stabroek News. (2017). *Erosion forces temporary closure of Shell Beach.* Accessed online [17/05/2021]. Available at: https://www.stabroeknews.com/2017/04/29/news/guyana/erosion-forces-temporary-closure-shell-beach/

Stanford Vision Lab. (2016). *ImageNet.* Accessed online [19/11/2019]. Available at http://www.image-net.org/.

Stevens, F.R., Gaughan, A.E., Linard, C. and Tatem, A.J. (2015). Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data. *PLOS ONE,* 10(2): 1-22.

Sun, L., Mi, X., Wei, J., Wang, J., Tian, X., Yu, H. and Gan, P. (2017). A cloud detection algorithm-generating method for remote sensing data at visible to short-wave infrared wavelengths. *International Society for Photogrammetry and Remote Sensing Journal of Photogrammetry and Remote Sensing,* 124(1): 70–88.

Suthaharan, S. 2016. A cognitive random forest: An intra-and intercognitive computing for big data classification under cune condition, in: *Handbook of statistics.* Elsevier, London, pp. 207-22.

Swift, D.J. (1974). Continental shelf sedimentation, in: *The geology of continental margins* Springer, Berlin, Heidelberg, pp. 117-135.

Tamiminia, H., Salehi, B., Mahdianpari, M., Quackenbush, L., Adeli, S. and Brisco, B. (2020). Google Earth Engine for geo-big data applications: A meta-analysis and systematic review. *ISPRS Journal of Photogrammetry and Remote Sensing,* 164(1): 152–170.

Tang, W., Zheng, M., Zhao, X., Shi, J., Yang, J. and Trettin, C.C. (2018). Big Geospatial Data Analytics for Global Mangrove Biomass and Carbon Estimation. *Sustainability,* 10(2): 472-489.

Tatem, A.J. (2017). WorldPop, open data for spatial demography. *Scientific Data,* 4(1): 1-4.

Thieler, E.R. and Danforth, W.W. (1994). Historical Shoreline Mapping (I): Improving Techniques and Reducing Positioning Errors. *Journal of Coastal Research,* 10(1): 549–563.

Thieler, E.R., Himmelstoss, E.A., Zichichi, J.L. and Ergul, A. (2009). *Digital Shoreline Analysis System (DSAS) version 4.0 – an ArcGIS extension for calculating shoreline change. US Geological Survey Open-file Report, 2008–1278. US Geological Survey, Reston, VA.* Accessed online [10/3/2020]. Available at: https://pubs.er.usgs.gov/publication/ofr20081278

Thieler, E.R., Smith, T.L., Knisel, J.M. and Sampson, D.W. (2012). *Massachusetts shoreline change mapping and analysis project, 2013 update. US Department of the Interior, US Geological Survey.* Accessed online [17/4/2019]. Available at: https://www.usgs.gov/centers/whcmsc/science/digital-shoreline-analysis-system-dsas?qt-science_center_objects=0#qt-science_center_objects

Thomas, C. (2020). Guyana and the advent of world-class petroleum finds, in: Looney, C. (2020). *Handbook of Caribbean Economies.* Routledge, U.K, pp. 289-303.

Thomas, N., Lucas, R., Bunting, P., Hardy, A., Rosenqvist, A. and Simard, M. (2017). Distribution and drivers of global mangrove forest change, 1996–2010. *PLOS ONE*, 12(6): 1-14.

Toms, B.A., Barnes, E.A. and Ebert-Uphoff, I. (2020). Physically Interpretable Neural Networks for the Geosciences: Applications to Earth System Variability. *Journal of Advances in Modeling Earth Systems*, 12(9): 1-20.

Toms, S. (2015). ArcPy and ArcGIS – Geospatial Analysis with Python. Packt Publishing Ltd, U.K.

Topouzelis, K. and Kitsiou, D. (2015). Detection and classification of mesoscale atmospheric phenomena above sea in SAR imagery. *Remote Sensing of Environment*, 160(1): 263–272.

Toure, S., Diop, O., Kpalma, K. and Maiga, A.S. (2019). Shoreline Detection using Optical Remote Sensing: A Review. *ISPRS International Journal of Geo-Information*, 8(2): 75-88.

Tsagkatakis, G., Aidini, A., Fotiadou, K., Giannopoulos, M., Pentari, A. and Tsakalides, P. (2019). Survey of Deep-Learning Approaches for Remote Sensing Observation Enhancement. *Sensors*, 19(18): 3929-3943.

Unberath, I., Vanierschot, L., Somers, B., Kerchove, R.V.D., Borre, J.V., Unberath, M. and Feilhauer, H. (2019). Remote sensing of coastal vegetation: Dealing with high species turnover by mapping multiple floristic gradients. *Applied Vegetation Science*, 22(4): 534–546.

United Nations. (2021). *Demographic and Social Statistics: World Population and Housing Census Programme- Census Documents.* Accessed online [29/06/2021]. Available at: https://unstats.un.org/unsd/demographic-social/census/document-resources/?search=&docType=Questionnaires&countryName=French+Guiana

University of East Anglia Climate Research Unit. (2021). *North Atlantic Oscillation.* Accessed online [15/05/2021] Available at: https://crudata.uea.ac.uk/cru/data/nao/values.htm

USGS. (2018*). Digital Shoreline Analysis System (DSAS). Version 5.0 User Guide.* Accessed online [15/04/2020]. Available at: https://www.usgs.gov/centers/whcmsc/science/digital-shoreline-analysis-system-dsas?qt-science_center_objects=0#qt-science_center_objects.

van Ledden, M., Vaughn, G., Lansen, J., Wiersma, F., Amsterdam, M. (2009). Extreme wave event along the Guyana coastline in October 2015. *Continental Shelf Research*, 29(1): 352–361.

Vaughn, S.E. (2017). Disappearing Mangroves: The Epistemic Politics of Climate Adaptation in Guyana. *Cultural Anthropology,* 32(2): 242–268.

Viavattene, C. (n.d.) *Coastal risk assessment framework guidance document*. Accessed online [05/10/2020]. Available at: http://eprints.mdx.ac.uk/18532/1/RISC-KIT_D2.3_CRAF_Guidance.pdf.

Vijay, V., Pimm, S.L., Jenkins, C.N. and Smith, S.J. (2016). The Impacts of Oil Palm on Recent Deforestation and Biodiversity Loss.*PLOS ONE,* 11(7): 1-19.

Voinov, A. and Gaddis, E.J.B. (2008). Lessons for successful participatory watershed modeling: A perspective from modelling practitioners. *Ecological Modelling, Special Issue dedicated to the memory of Yuri Svirezhev*, 216(1): 197–207.

Vos, K., Harley, M.D., Splinter, K.D., Simmons, J.A. and Turner, I.L. (2019a). Sub-annual to multi-decadal shoreline variability from publicly available satellite imagery. *Coastal Engineering*, 150(1): 160-174.

Vos, K., Splinter, K.D., Harley, M.D., Simmons, J.A. and Turner, I.L. (2019b). CoastSat: A Google Earth Engine-enabled Python toolkit to extract shorelines from publicly available satellite imagery. *Environmental Modelling & Software*, 122(1): 1-7.

Vousdoukas, M.I., Ranasinghe, R., Mentaschi, L., Plomaritis, T.A., Athanasiou, P., Luijendijk, A. and Feyen, L. (2020). Sandy coastlines under threat of erosion. *Nature Climate Chang*e, 10(3): 260–263.

Wadey, M.P., Haigh, I.D., Nicholls, R.J., Brown, J.M., Horsburgh, K., Carroll, B., Gallop, S.L., Mason, T. and Bradshaw, E. (2015). A comparison of the 31 January – 1 February 1953 and 5 – 6 December 2013 coastal flood events around the UK. *Frontiers in Marine Science,* 2(1): 84-111.

Wagner, W., Lague, D., Mohrig, D., Passalacqua, P., Shaw, J. and Moffett, K. (2017). Elevation change and stability on a prograding delta. *Geophysical Research Letters*, 44(4): 1786-1794.

Walcker, R., Anthony, E.J., Cassou, C., Aller, R.C., Gardel, A., Proisy, C., Martinez, J.-M. and Fromard, F. (2015). Fluctuations in the extent of mangroves driven by multi-decadal changes in North Atlantic waves. *Journal of Biogeography*, 42(11): 2209–2219.

Waldner, F. and Diakogiannis, F.I. (2020). Deep learning on edge: Extracting field boundaries from satellite images with a convolutional neural network. *Remote Sensing of Environment*, 245(1): 1-14.

Wan, L., Zhang, H., Lin, G. and Lin, H. (2019). A small-patched convolutional neural network for mangrove mapping at species level using high-resolution remote-sensing image. *Annals of GIS*, 25(1): 45-55.

Wang, L., Jia, M., Yin, D. and Tian, J. (2019). A review of remote sensing for mangrove forests: 1956–2018. *Remote Sensing of Environment*, 231(1): 1-23.

Wang, Y. (2009). Remote Sensing of Coastal Environments. CRC Press, London.

Wang, L., Ting, M. and Kushner, P.J. (2017). A robust empirical seasonal prediction of winter NAO and surface climate. *Scientific reports*, 7(1): 1-9.

Wang, Z., Lam, N.S.N., Obradovich, N. and Ye, X. (2019). Are vulnerable communities digitally left behind in social responses to natural disasters? An evidence from Hurricane Sandy with Twitter data. *Applied Geography*, 108(1): 1–8.

Watkins, B. and Van Niekerk, A. (2019). A comparison of object-based image analysis approaches for field boundary delineation using multi-temporal Sentinel-2 imagery. *Computers and Electronics in Agriculture*, 158(1): 294-302.

Weinstein, B.G., Marconi, S., Bohlman, S.A., Zare, A., Singh, A., Graves, S.J. and White, E.P. (2021). A remote sensing derived data set of 100 million individual tree crowns for the National Ecological Observatory Network. *eLife*, 1(1): 1-18.

Wen, L. and Hughes, M. (2020). Coastal wetland mapping using ensemble learning algorithms: A comparative study of bagging, boosting and stacking techniques. *Remote Sensing*, 12(10): 1683-1700.

Whitehead, N.L. and Simon, G. (2010). Materializing the Past among the Lokono (Arawak) of the Berbice River, Guyana. *Antropológica*, 54(114): 87-127

Wilson, K., Lentz, E.E., Miselis, J.L., Safak, I. and Brenner, O.T. (2019). A Bayesian Approach to Predict Sub-Annual Beach Change and Recovery. *Estuaries and Coasts*, 42(1): 112–131.

Wilson, K.E., Adams, P.N., Hapke, C.J., Lentz, E.E. and Brenner, O. (2015). Application of Bayesian Networks to hindcast barrier island morphodynamics. *Coastal Engineering,* 102(2): 30–43.

Winterwerp, J.C., Erftemeijer, P.L.A., Suryadiputra, N., van Eijk, P. and Zhang, L. (2013). Defining Eco-Morphodynamic Requirements for Rehabilitating Eroding Mangrove-Mud Coasts. *Wetlands,* 33(3): 515–526.

Winterwerp, J.C., Graaff, R.F. de, Groeneweg, J., Luijendijk, A.P.(2007). *Modelling of wave damping at Guyana mud coast. Coastal Engineering 54, 249–261. World Bank, 2018.Suriname Coastal Resilience assessment. Technical Report 130439.* Accessed Online [25/06/2021]. Available online: https://documents1.worldbank.org/curated/en/684611538551863364/pdf/Suriname-Coastal-Resilience-Assessment-Feb-9-Low-Res.pdf

World Bank. (2018). *GDP per capital (current US$).* Accessed Online [01/06/2021]. Available at: https://data.worldbank.org/indicator/NY.GDP.PCAP.CD

WorldPop. (2021). *Population Counts: Population Counts/ Unconstrained individual countries 2000 – 2020 UN adjusted (100 m resolution).* Accessed online [01/02/2021]. Available at: https://www.worldpop.org/geodata/listing?id=69

Wright, L.D. and Short, A.D. (1984). Morphodynamic variability of surf zones and beaches: a synthesis. *Marine Geology,* 56(1-4): 93-118.

Wu, K., Li, X.-M. and Huang, B.(2021). Retrieval of Ocean Wave Heights From Spaceborne SAR in the Arctic Ocean With a Neural Network. *Journal of Geophysical Research: Oceans,* 126(3): 1-11.

Xie, S. and Tu, Z. (2015). Holistically-nested edge detection. *Proceedings of the IEEE International Conference on Computer Vision*, 1(1): 1395-1403.

Yang, H., Li, L., Hu, H., Wu, Y., Xia, H., Liu, Y. and Tan, S. (2018). A coastline generalization method that considers buffer consistency. *PLOS ONE,* 13(11): 1-16.

Yosinski, J., Clune, J., Bengio, Y. and Lipson, H. (2014). How transferable are features in deep neural networks? *Machine Learning*, 1411(1): 1-8.

Young, I.R. and Ribal, A. (2019). Multiplatform evaluation of global trends in wind speed and wave height. *Science,* 364(6440): 548–552.

Young, I.R., Zieger, S. and Babanin, A.V. (2011). Global Trends in Wind Speed and Wave Height. *Science,* 332(6028): 451–455.

Young, N.E., Anderson, R.S., Chignell, S.M., Vorster, A.G., Lawrence, R. and Evangelista, P.H. (2017). A survival guide to Landsat preprocessing. *Ecology,* 98(4): 920–932.

Yuan, S., Luo, X., Mu, B., Li, J. and Dai, G. (2019). Prediction of North Atlantic Oscillation index with convolutional LSTM based on ensemble empirical mode decomposition. *Atmosphere*, 10(5): 252-263.

Yuan, F. and Liu, R., (2020). Identifying damage-related social media data during Hurricane Matthew: A machine learning approach, in: Cho, Y.K., Leite, F., Behzadan, A. and Wang, C. (Eds.), (2019). *Computing in Civil Engineering 2019: Visualization, Information Modeling, and Simulation.* American Society of Civil Engineers, Reston, VA, pp. 207-214.

Zarillo, G.A., Kelley, J. and Larson, V. (2008). *A GIS Based Tool for Extracting Shoreline Positions from Aerial Imagery (BeachTools) Revised (No. ERDC/CHL-CHETN-IV-73). Engineer research and development center Vicksburg MS coastal and hydraulics lab.* Accessed online [3/11/2019]. Available at: https://apps.dtic.mil/sti/pdfs/ADA490237.pdf

Zhang, H., Jiang, Q. and Xu, J. (2013). Coastline Extraction Using Support Vector Machine from Remote Sensing Image. *Journal of Multimedia*, 8(2):175-182.

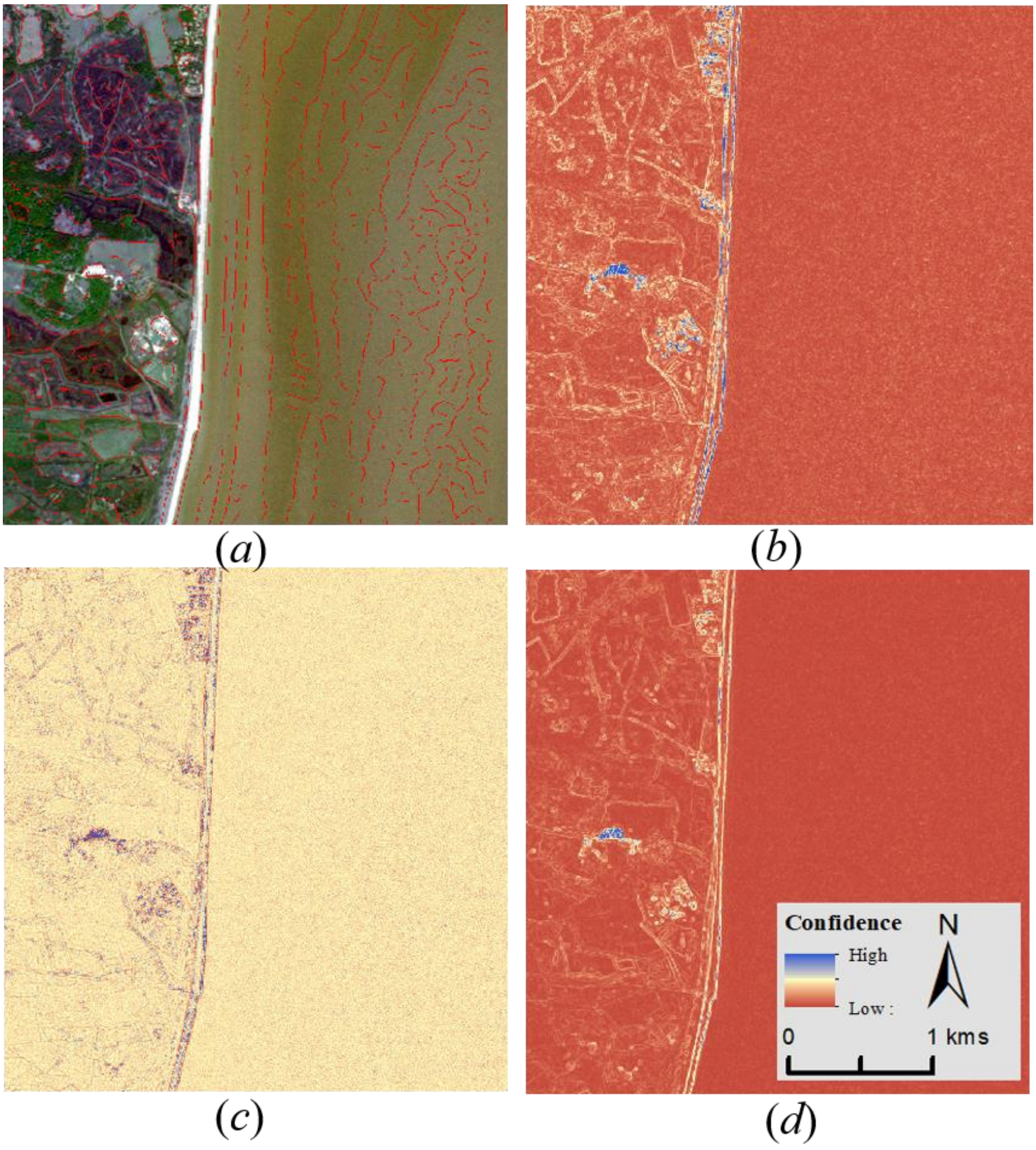
Zhang, L., Zhang, L. and Du, B. (2016). Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art. *IEEE Geoscience and Remote Sensing Magazine,* 4(2): 22–40.

Zhao, B., Guo, H., Yan, Y., Wang, Q. and Li, B. (2008). A simple waterline approach for tidelands using multi-temporal satellite images: A case study in the Yangtze Delta. *Estuarine, Coastal and Shelf Science,* 77(1): 134–142.

Zhu, Z., Wang, S. and Woodcock, C.E. (2015). Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sensing of Environment*, 159(1): 269–277.

Supplemental Materials A

1. Edge detection operators using greyscale images



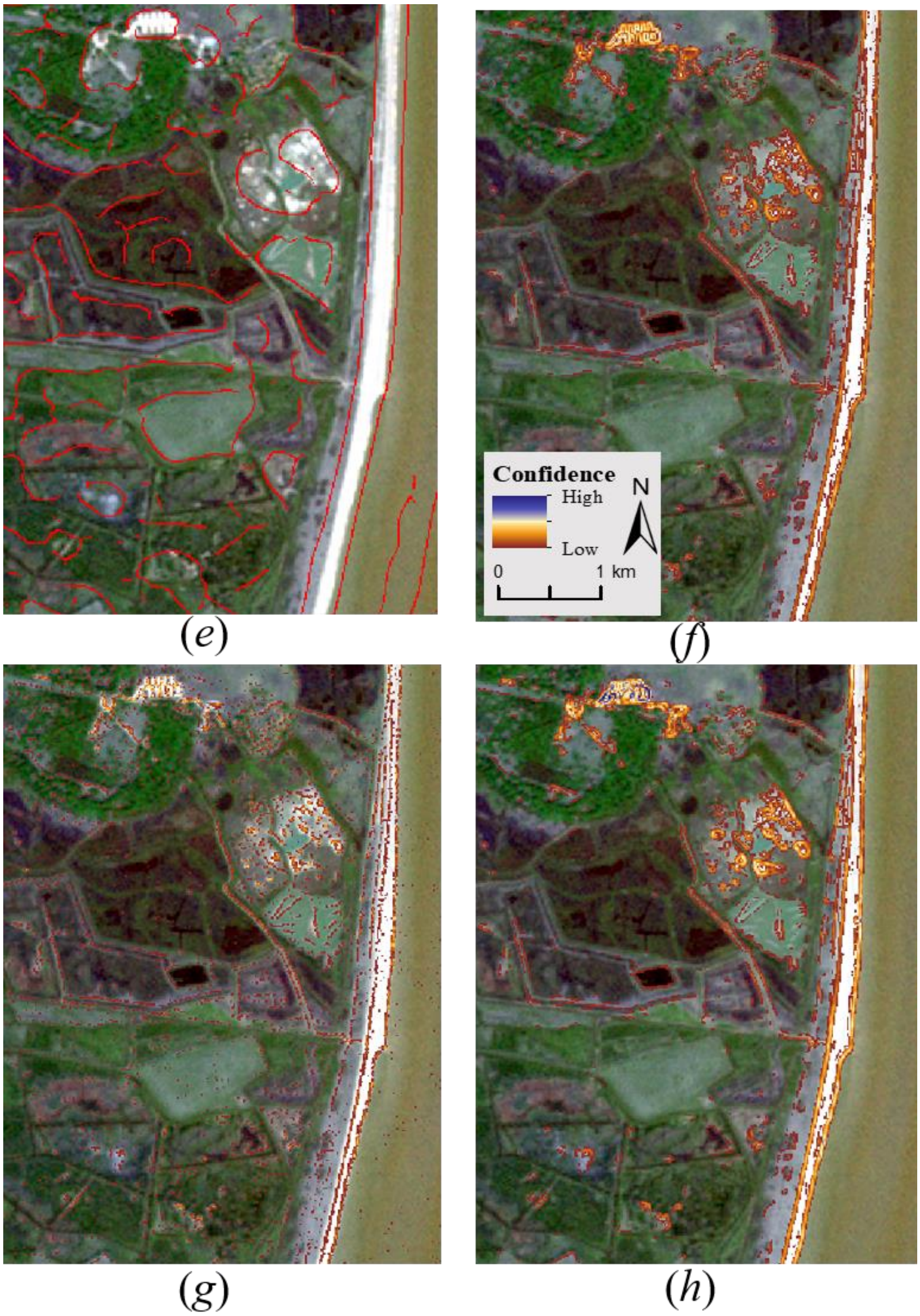
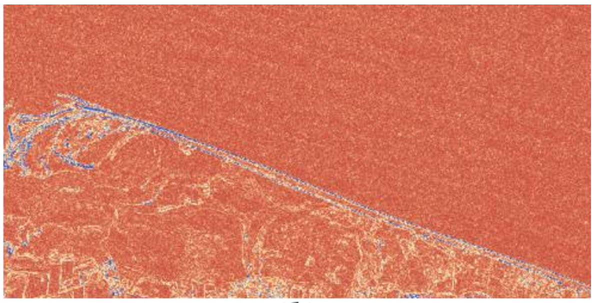


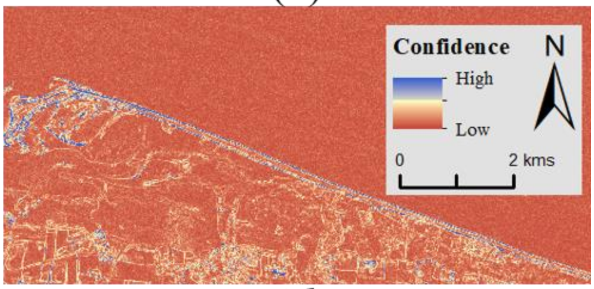
Figure A: Filtered and unfiltered outputs from applying edge detection operators to the greyscale image of Dunwich, Suffolk, UK. Unfiltered outputs from applying (*a*) Canny edge, (*b*) Roberts, (*c*) Laplacian and (*d*) Sobel edge detection to the greyscale image. Filtered outputs to only show the 5% of pixels with the highest values after applying (*e*) Canny edge, (*f*) Roberts, (*g*) Laplacian and (*h*) Sobel edge detection to the greyscale image. Canny edge outputs in (*a*) and (*e*) are binary images where pixels predicted to be an edge are shown in red. For all other images, blue and red correspond to pixels determined to be an edge with high and low confidence respectively.

(*a*)

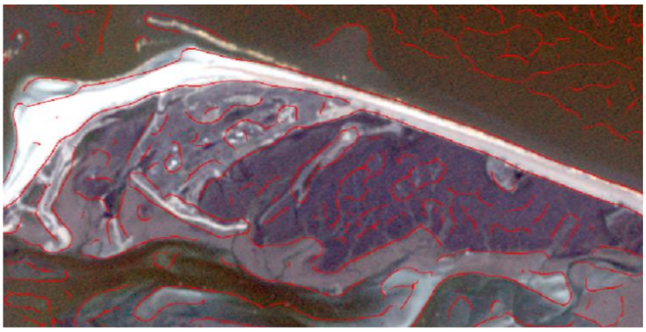


(*b*)

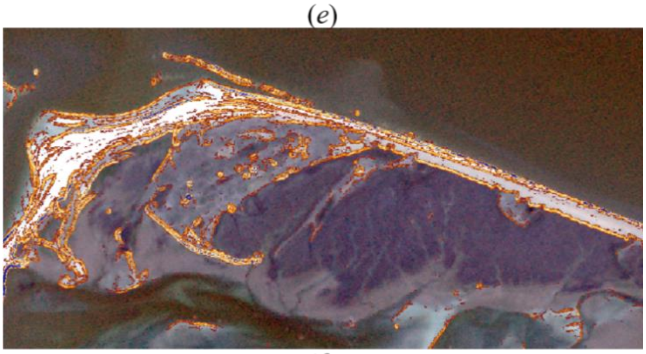
(*c*)



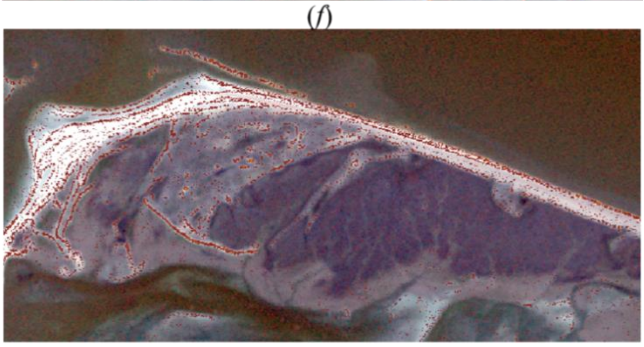
(*d*)



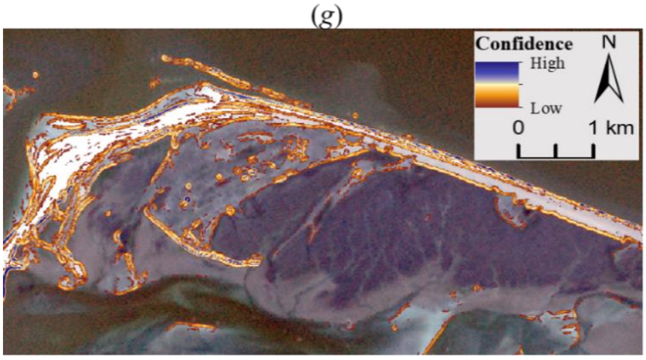
(*e*)



(*f*)

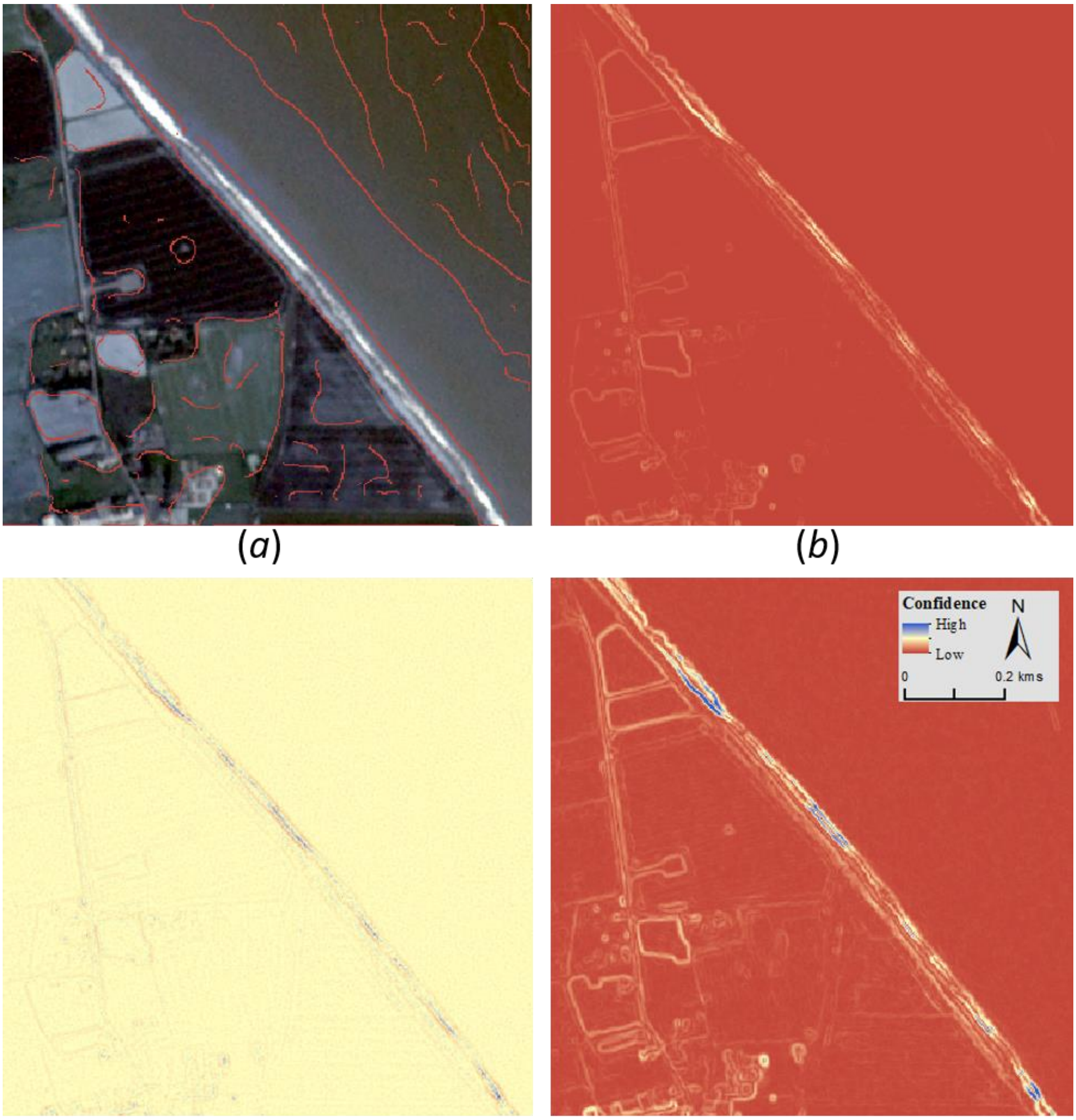


(*g*)



(*h*)

Figure B: Filtered and unfiltered outputs from applying edge detection operators to the greyscale image of Blakeney Point, Norfolk, UK. Unfiltered outputs from applying (*a*) Canny edge, (*b*) Roberts, (*c*) Laplacian and (*d*) Sobel edge detection to the greyscale image. Filtered outputs to only show the 5% of pixels with the highest values after applying (*e*) Canny edge, (*f*) Roberts, (*g*) Laplacian and (*h*) Sobel edge detection to the greyscale image. Canny edge outputs in (*a*) and (*e*) are binary images where pixels predicted to be an edge are shown in red. For all other images, blue and red correspond to pixels determined to be an edge with high and low confidence respectively.



(*a*) (*b*)

(*c*) (*d*)

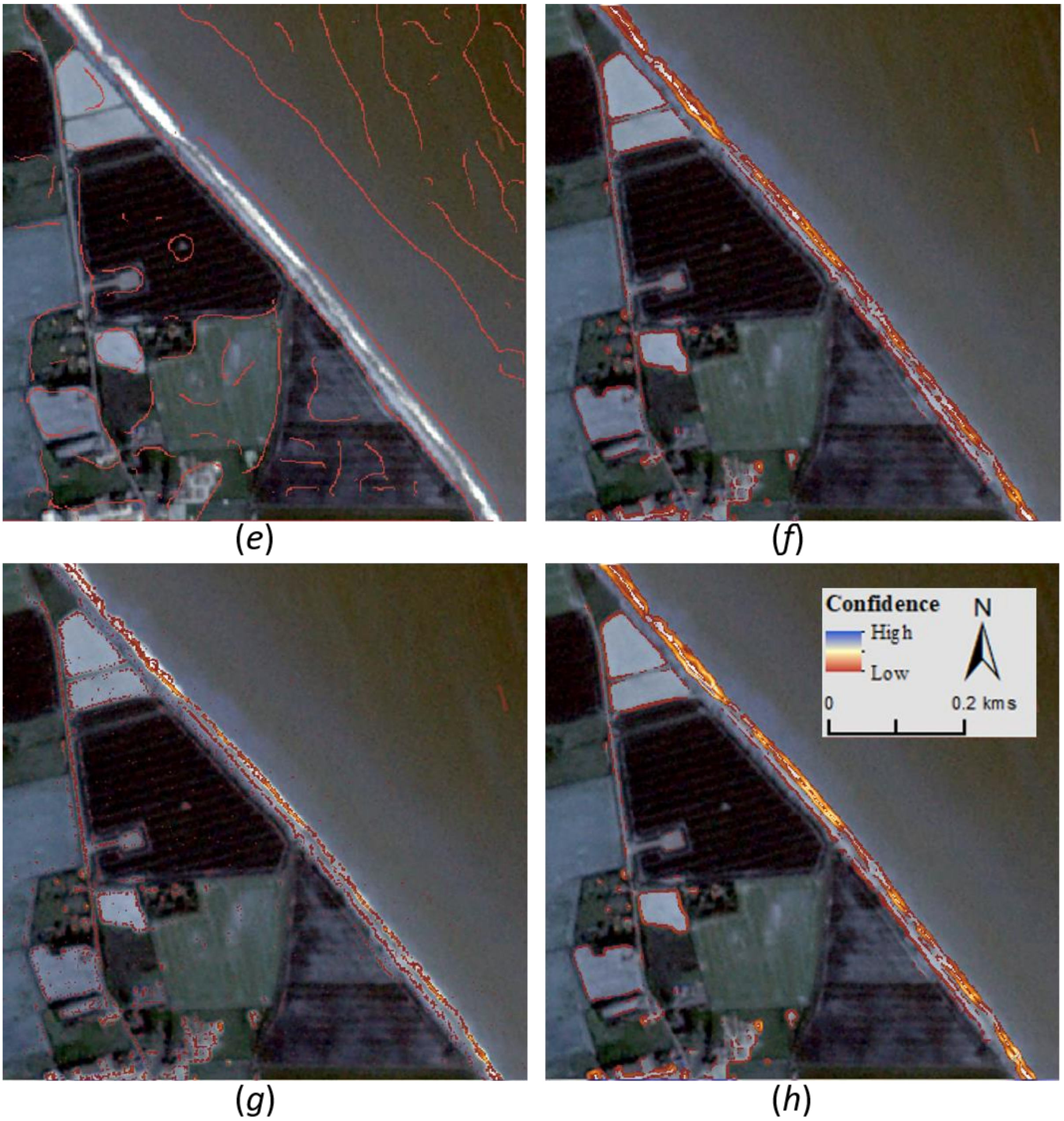


Figure C: Filtered and unfiltered outputs from applying edge detection operators to the greyscale image of Holderness, East Yorkshire, UK. Unfiltered outputs from applying (*a*) Canny edge, (*b*) Roberts, (*c*) Laplacian and (*d*) Sobel edge detection to the greyscale image. Filtered outputs to only show the 5% of pixels with the highest values after applying (*e*) Canny edge, (*f*) Roberts, (*g*) Laplacian and (*h*) Sobel edge detection to the greyscale image. Canny edge outputs in (*a*) and (*e*) are binary images where pixels predicted to be an edge are shown in red. For all other images, blue and red correspond to pixels determined to be an edge with high and low confidence respectively.

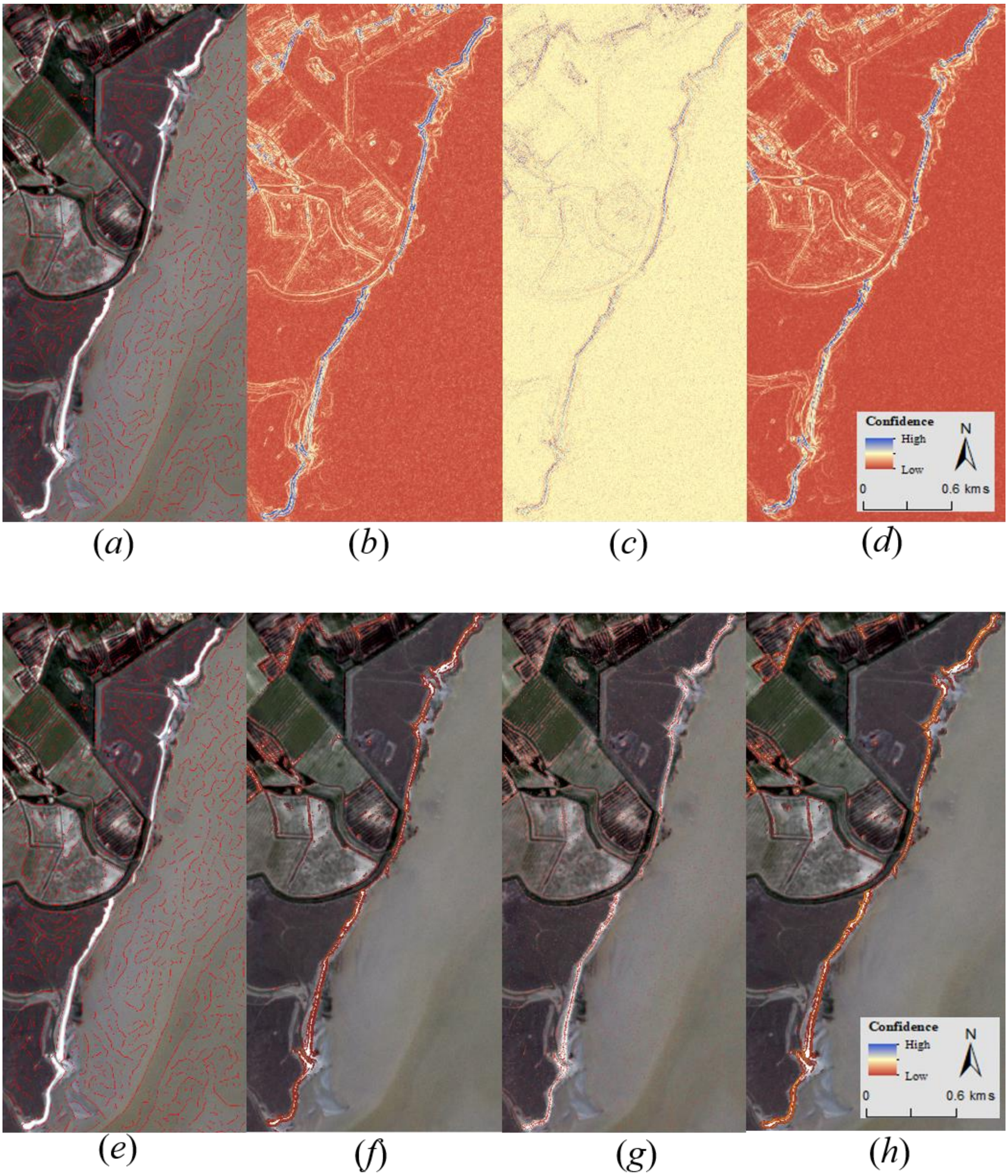


Figure D: Filtered and unfiltered outputs from applying edge detection operators to the greyscale image of Hornsey, Essex, UK. Unfiltered outputs from applying (*a*) Canny edge, (*b*) Roberts, (*c*) Laplacian and (*d*) Sobel edge detection to the greyscale image. Filtered outputs to only show the 5% of pixels with the highest values after applying (*e*) Canny edge, (*f*) Roberts, (*g*) Laplacian and (*h*) Sobel edge detection to the greyscale image. Canny edge outputs in (*a*) and (*e*) are binary images where pixels predicted to be an edge are shown in red. For all other images, blue and red correspond to pixels determined to be an edge with high and low confidence respectively.

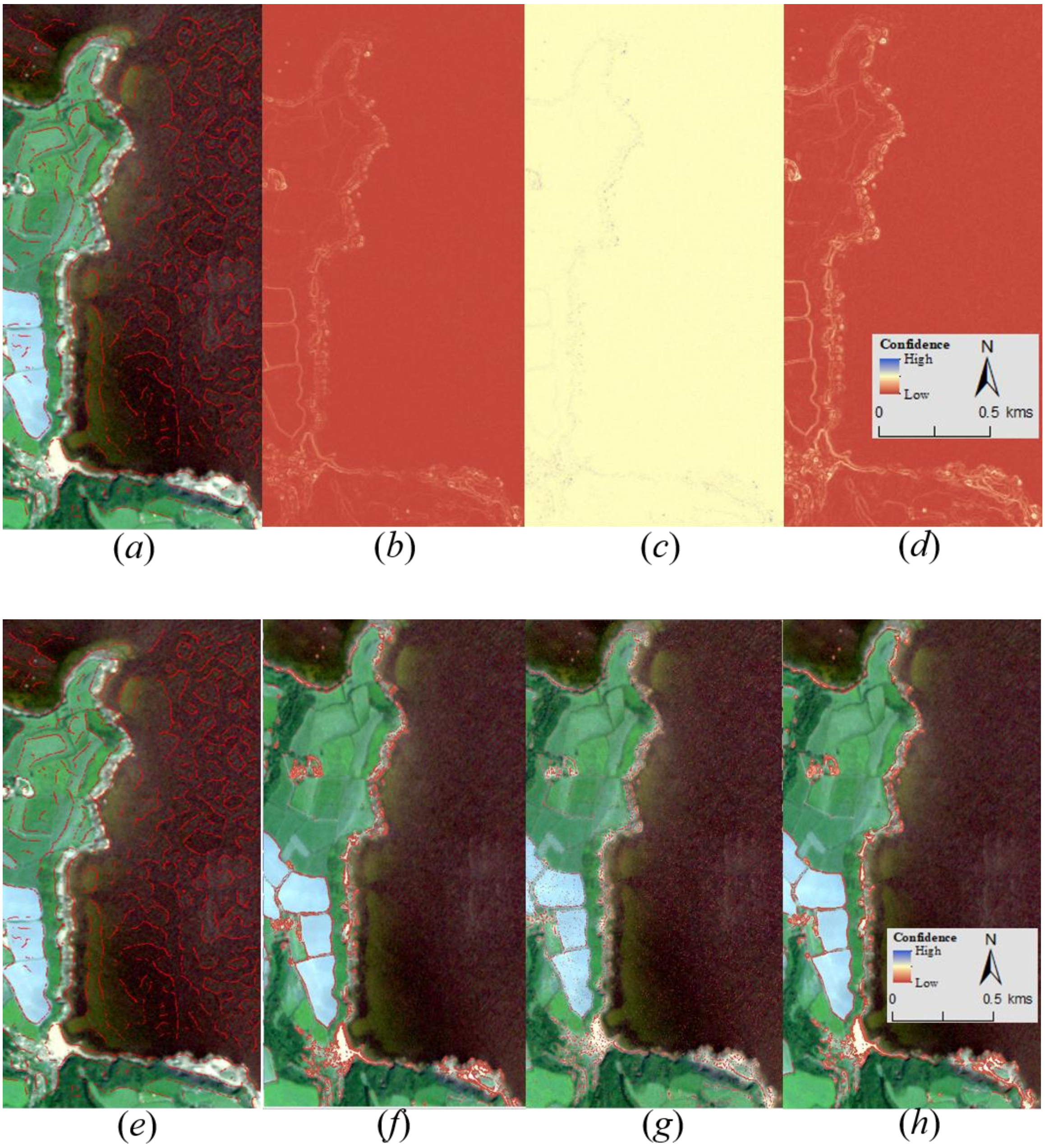
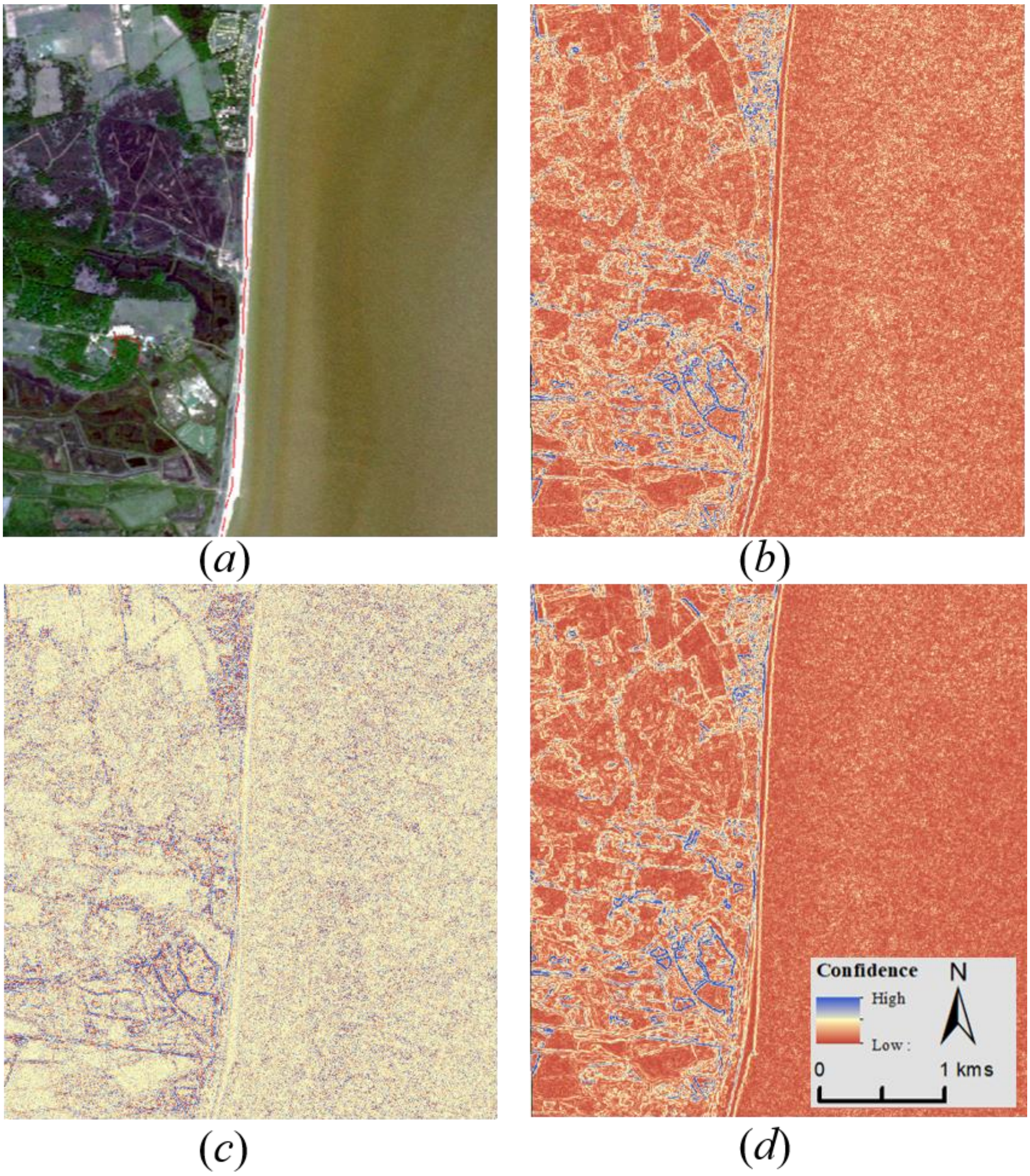


Figure E: Filtered and unfiltered outputs from applying edge detection operators to the greyscale image of Porthallow, Cornwall, UK. Unfiltered outputs from applying (*a*) Canny edge, (*b*) Roberts, (*c*) Laplacian and (*d*) Sobel edge detection to the greyscale image. Filtered outputs to only show the 5% of pixels with the highest values after applying (*e*) Canny edge, (*f*) Roberts, (*g*) Laplacian and (*h*) Sobel edge detection to the greyscale image. Canny edge outputs in (*a*) and (*e*) are binary images where pixels predicted to be an edge are shown in red. For all other images, blue and red correspond to pixels determined to be an edge with high and low confidence respectively.

2. Edge Detection operator outputs using NDVI image



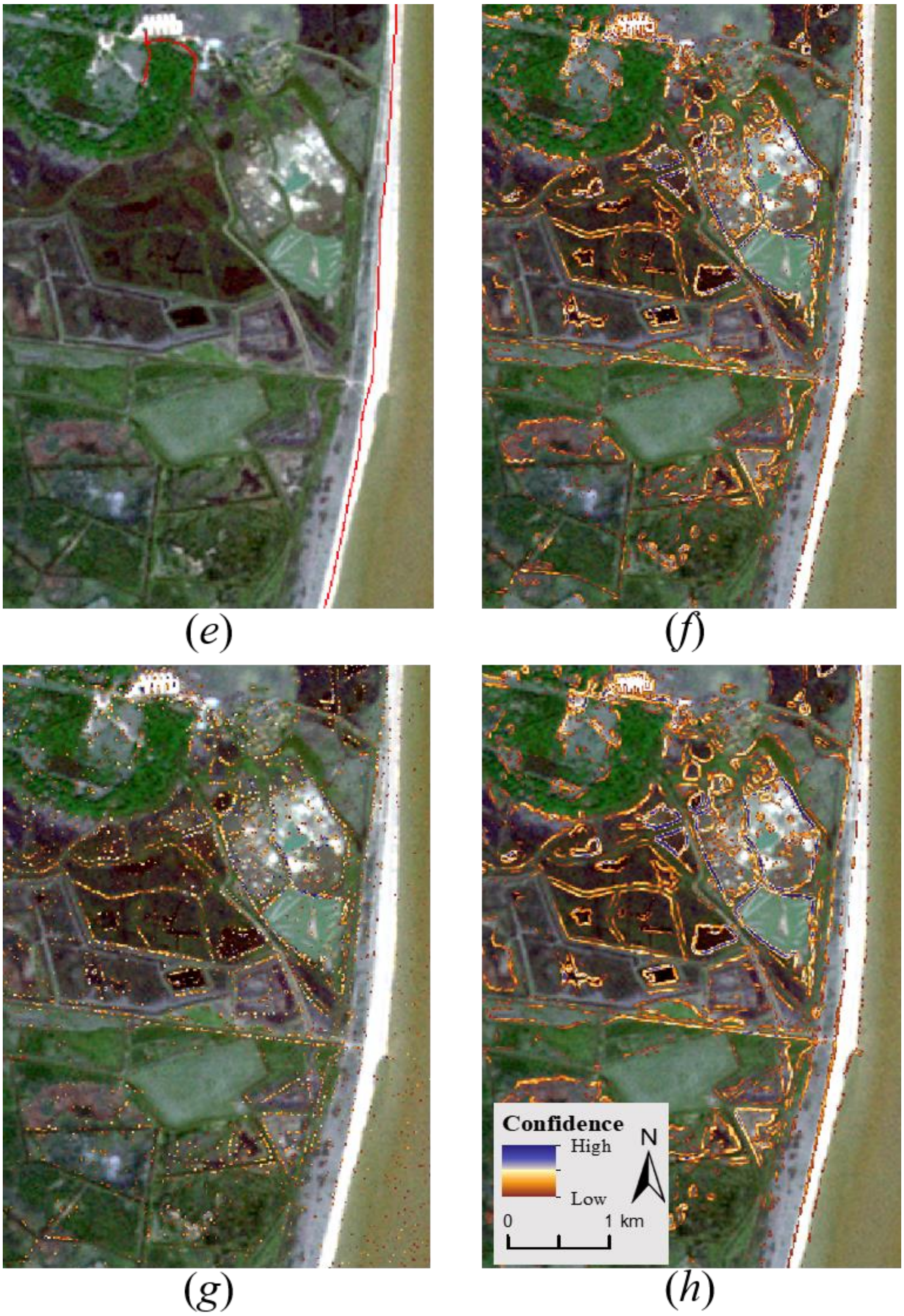
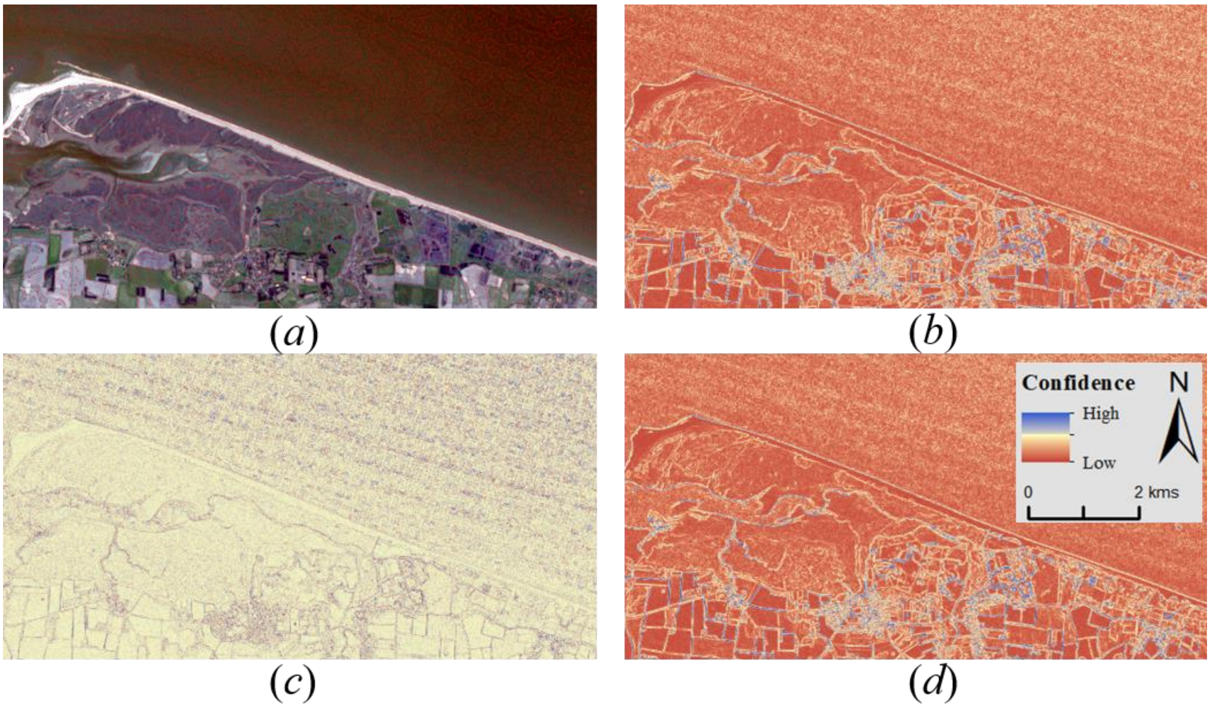
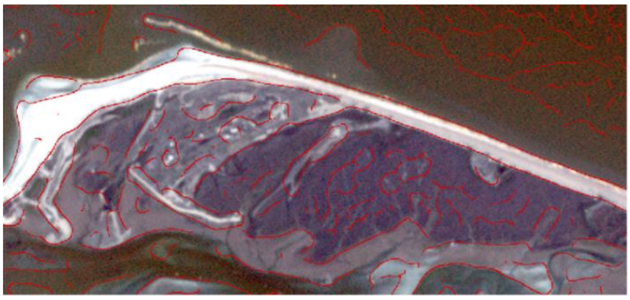
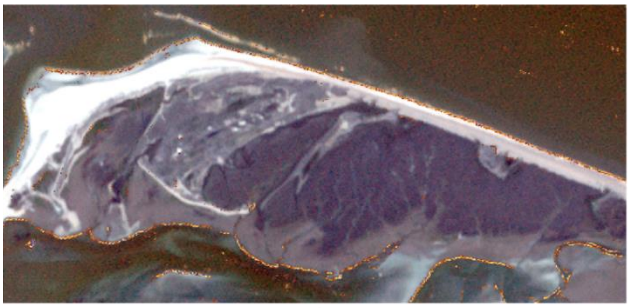


Figure F: Filtered and unfiltered outputs from applying edge detection operators to the NDVI image of Dunwich, Suffolk, UK. Unfiltered outputs from applying (*a*) Canny edge, (*b*) Roberts, (*c*) Laplacian and (*d*) Sobel edge detection to the greyscale image. Filtered outputs to only show the 5% of pixels with the highest values after applying (*e*) Canny edge, (*f*) Roberts, (*g*) Laplacian and (*h*) Sobel edge detection to the greyscale image. Canny edge outputs in (*a*) and (*e*) are binary images where pixels predicted to be an edge are shown in red. For all other images, blue and red correspond to pixels determined to be an edge with high and low confidence respectively.

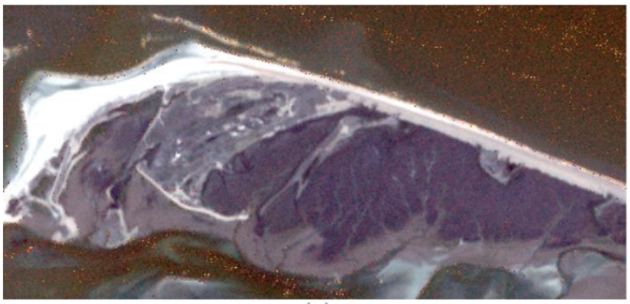




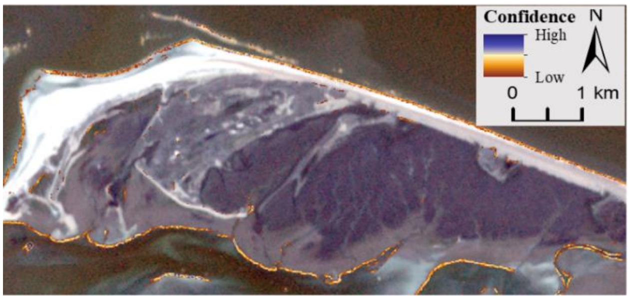
(*e*)



(*f*)

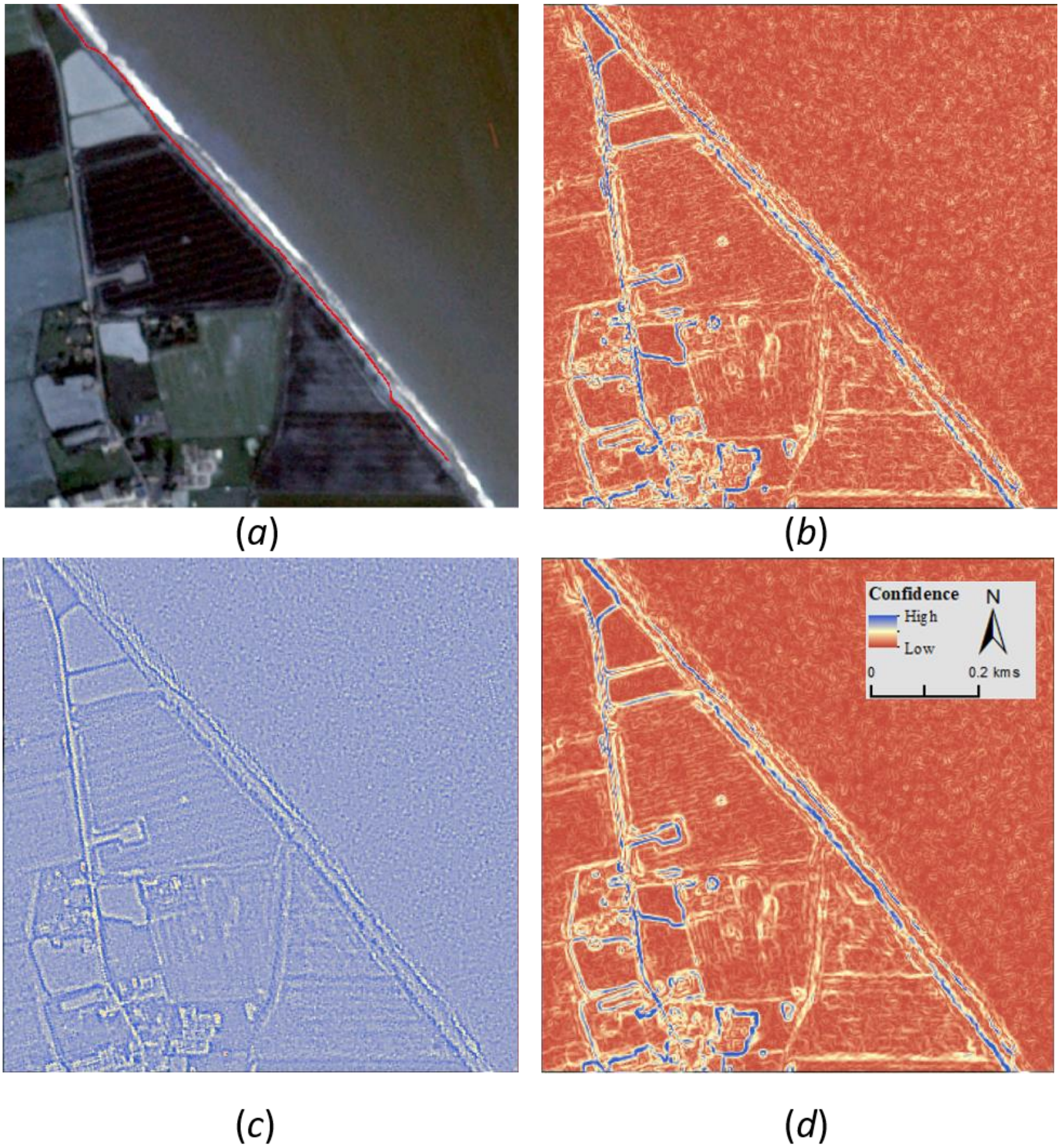


(*g*)



(*h*)

Figure G: Filtered and unfiltered outputs from applying edge detection operators to the NDVI image of Blakeney, Norfolk, UK. Unfiltered outputs from applying (*a*) Canny edge, (*b*) Roberts, (*c*) Laplacian and (*d*) Sobel edge detection to the greyscale image. Filtered outputs to only show the 5% of pixels with the highest values after applying (*e*) Canny edge, (*f*) Roberts, (*g*) Laplacian and (*h*) Sobel edge detection to the greyscale image. Canny edge outputs in (*a*) and (*e*) are binary images where pixels predicted to be an edge are shown in red. For all other images, blue and red correspond to pixels determined to be an edge with high and low confidence respectively.


(a)
(b)
Confidence
High
Low
N
0
0.2 kms
(c)
(d)

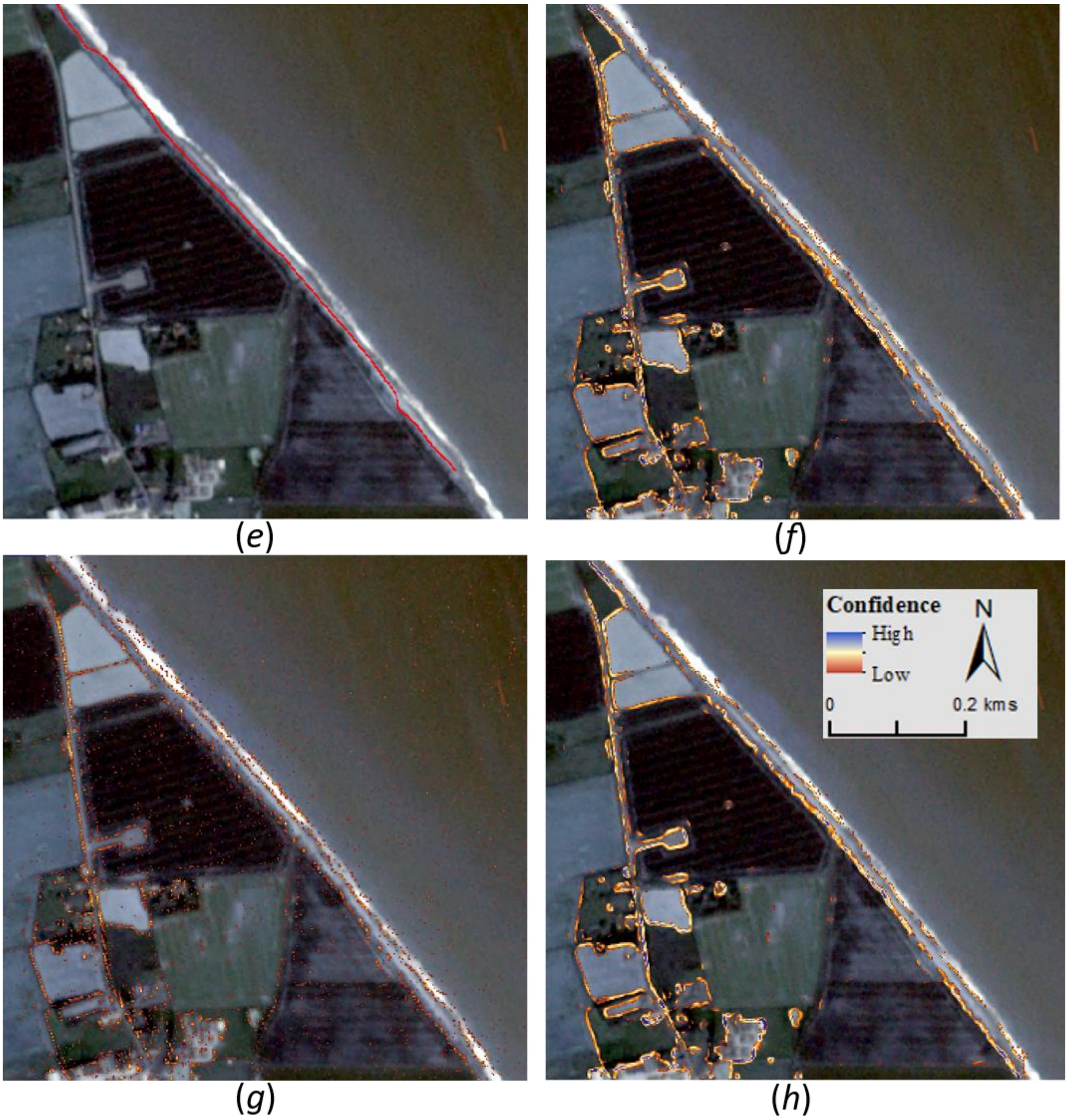


Figure H: Filtered and unfiltered outputs from applying edge detection operators to the NDVI image of Holderness, East Yorkshire, UK. Unfiltered outputs from applying (*a*) Canny edge, (*b*) Roberts, (*c*) Laplacian and (*d*) Sobel edge detection to the greyscale image. Filtered outputs to only show the 5% of pixels with the highest values after applying (*e*) Canny edge, (*f*) Roberts, (*g*) Laplacian and (*h*) Sobel edge detection to the greyscale image. Canny edge outputs in (*a*) and (*e*) are binary images where pixels predicted to be an edge are shown in red. For all other images, blue and red correspond to pixels determined to be an edge with high and low confidence respectively.

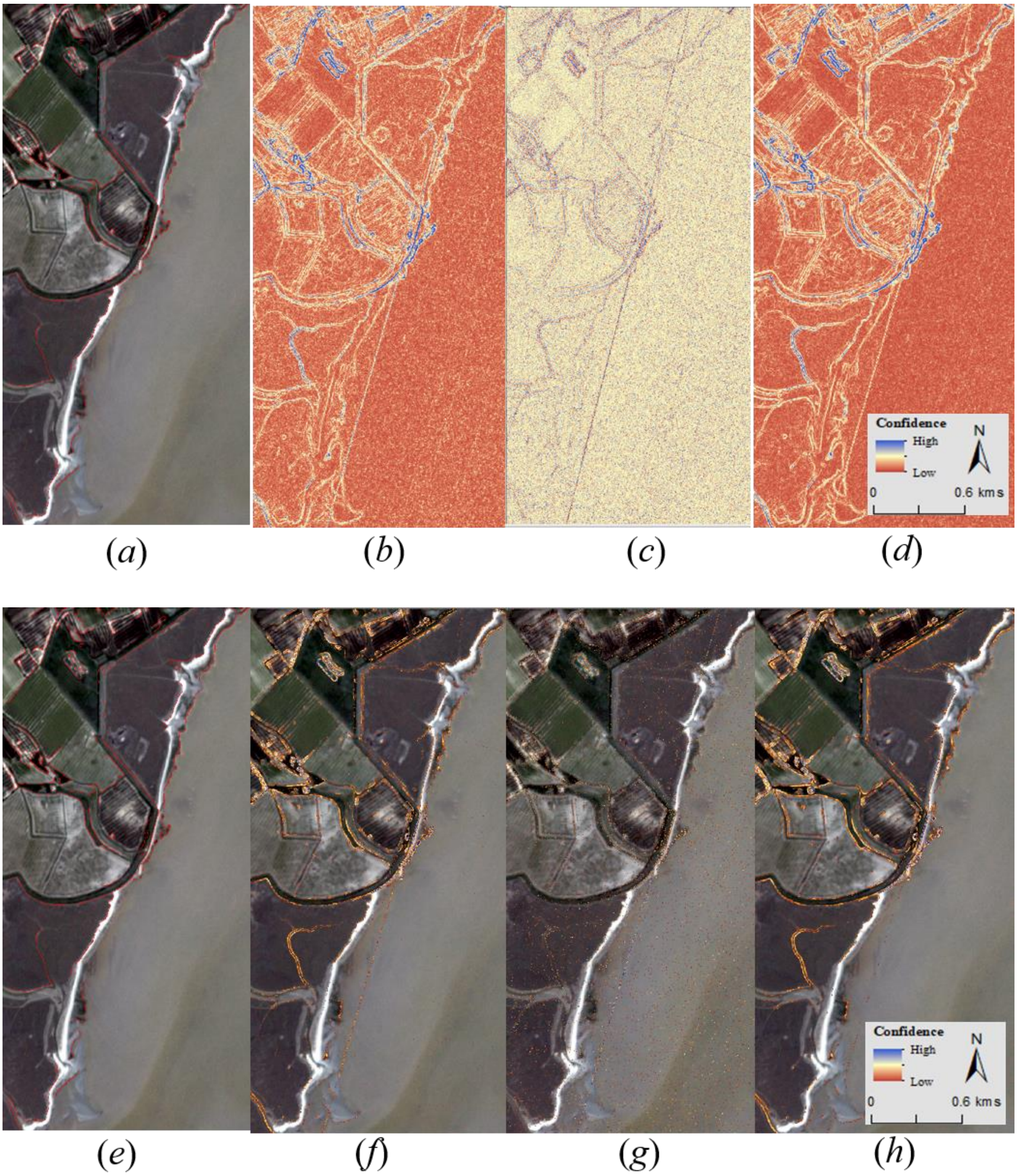
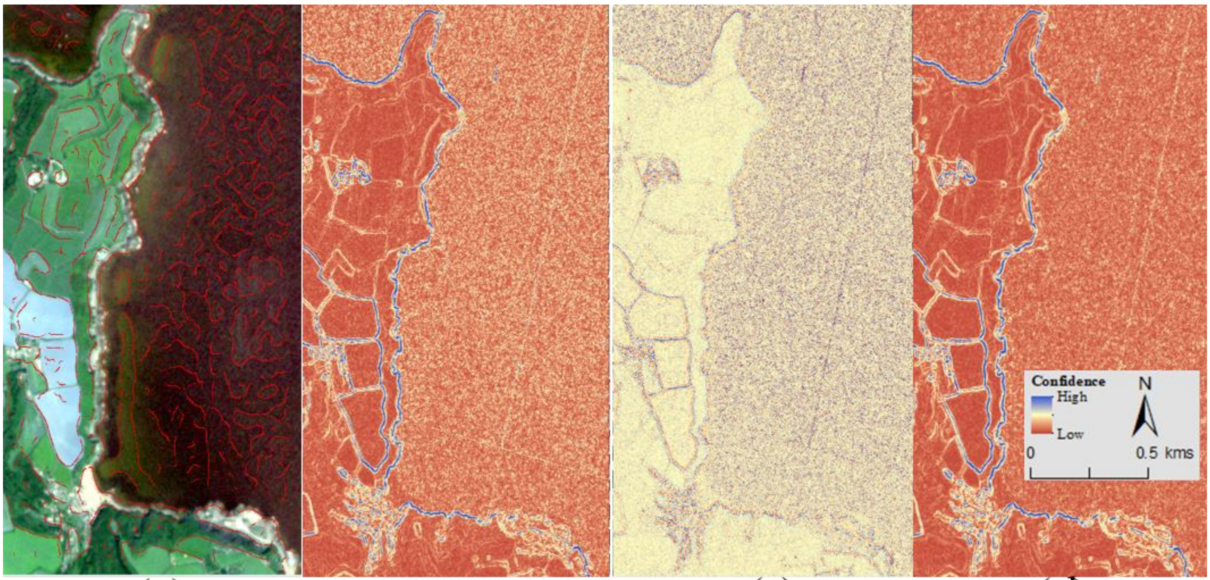


Figure I: Filtered and unfiltered outputs from applying edge detection operators to the NDVI image of Hornsey, Essex, UK. Unfiltered outputs from applying (*a*) Canny edge, (*b*) Roberts, (*c*) Laplacian and (*d*) Sobel edge detection to the greyscale image. Filtered outputs to only show the 5% of pixels with the highest values after applying (*e*) Canny edge, (*f*) Roberts, (*g*) Laplacian and (*h*) Sobel edge detection to the greyscale image. Canny edge outputs in (*a*) and (*e*) are binary images where pixels predicted to be an edge are shown in red. For all other images, blue and red correspond to pixels determined to be an edge with high and low confidence respectively.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure J: Filtered and unfiltered outputs from applying edge detection operators to the NDVI image of Porthallow, Cornwall, UK. Unfiltered outputs from applying (*a*) Canny edge, (*b*) Roberts, (*c*) Laplacian and (*d*) Sobel edge detection to the greyscale image. Filtered outputs to only show the 5% of pixels with the highest values after applying (*e*) Canny edge, (*f*) Roberts, (*g*) Laplacian and (*h*) Sobel edge detection to the greyscale image. Canny edge outputs in (*a*) and (*e*) are binary images where pixels predicted to be an edge are shown in red. For all other images, blue and red correspond to pixels determined to be an edge with high and low confidence respectively.

3. Support Vector machines outputs

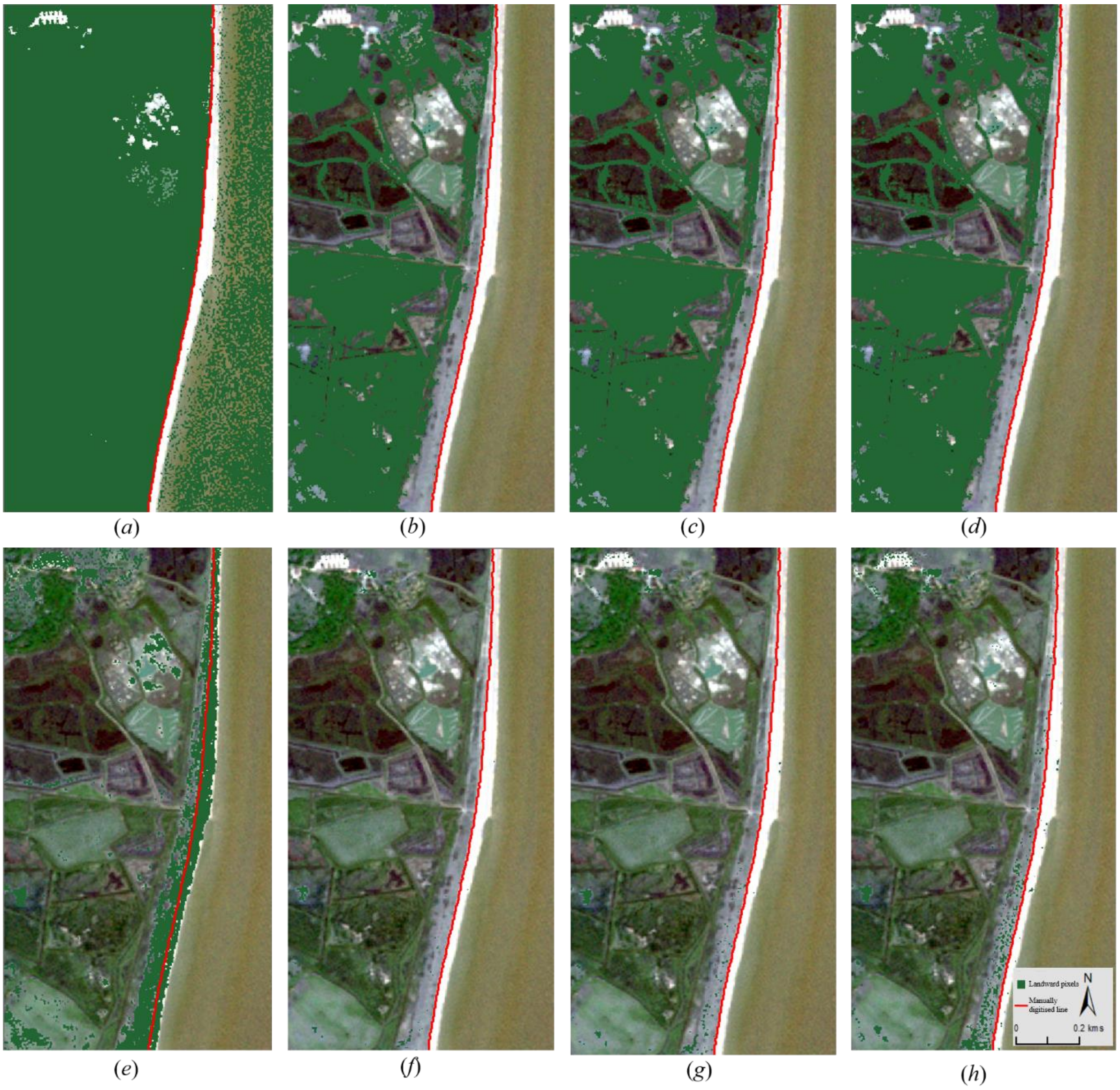
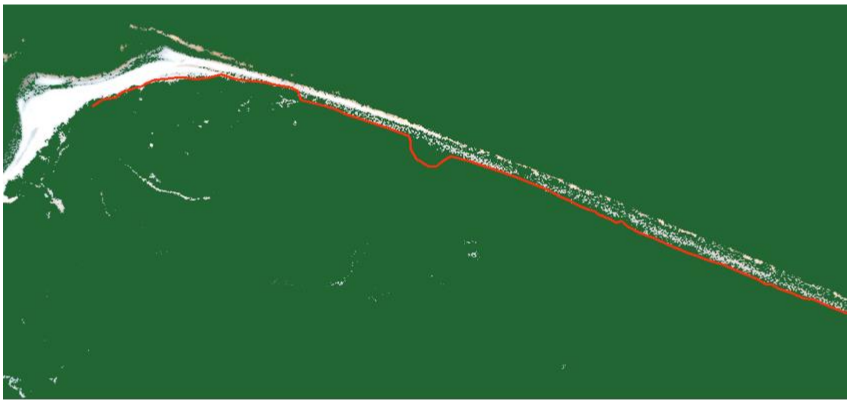


Figure K: Support Vector Machines (SVM) outputs from Dunwich, Suffolk, UK. SVM output using a linear model and regularisation parameter value, C, of (*a*) 1, (*b*) 2, (*c*) 3 and (*d*) 4. SVM output using a polynomial kernel and a regularisation parameter value, C, of (*e*) 1, (*f*) 2, (*g*) 3 and (*h*) 4. For all images, the pixels in green correspond to pixels predicted by the SVM to be positioned landwards of the coastal vegetation line. The manually digitised line is in red.



(*a*)

(*b*)

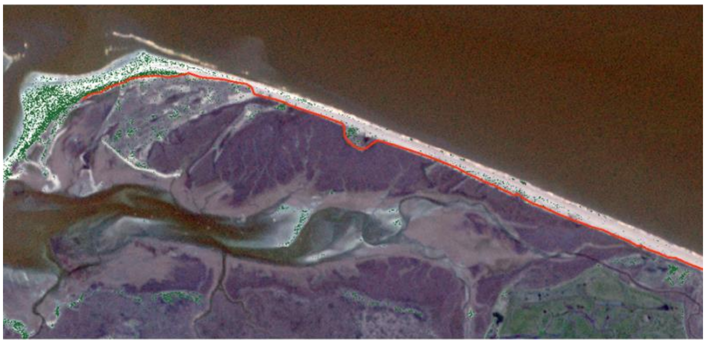
(*c*)

(*d*)

(*e*)

(*f*)

(*g*)



(*h*)

Figure L: Support Vector Machines (SVM) outputs from Blakeney, Norfolk, UK. SVM output using a linear model and regularisation parameter value, C, of (*a*) 1, (*b*) 2, (*c*) 3 and (*d*) 4. SVM output using a polynomial kernel and a regularisation parameter value, C, of (*e*) 1, (*f*) 2, (*g*) 3 and (*h*) 4. For all images, the pixels in green correspond to pixels predicted by the SVM to be positioned landwards of the coastal vegetation line. The manually digitised line is in red.

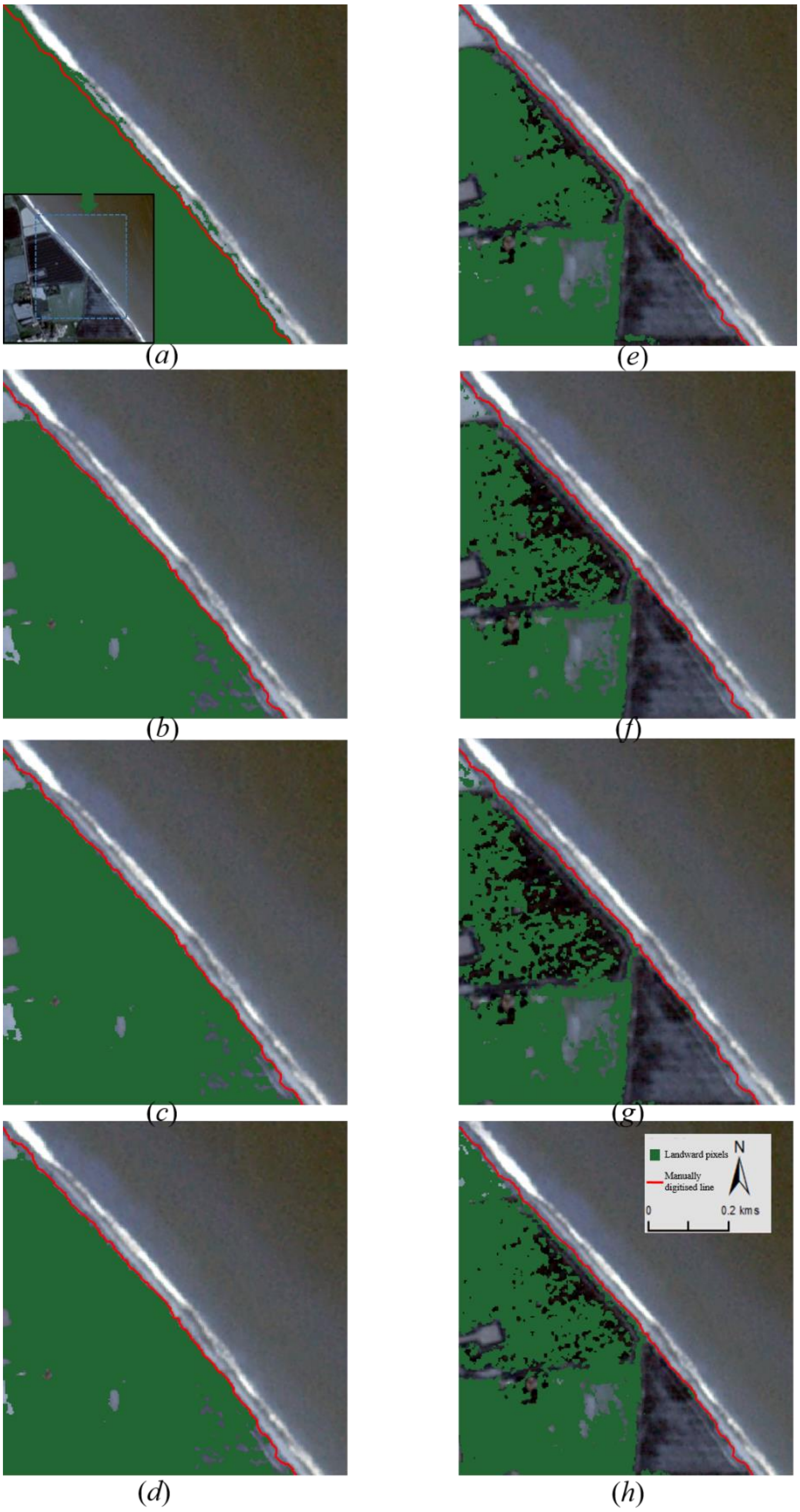


Figure M: Support Vector Machines (SVM) outputs from Holderness, East Yorkshire, UK. SVM output using a linear model and regularisation parameter value, C, of (*a*) 1, (*b*) 2, (*c*) 3 and (*d*) 4. SVM output using a polynomial kernel and a regularisation parameter value, C, of (*e*) 1, (*f*) 2, (*g*) 3 and (*h*) 4. For all images, the pixels in green correspond to pixels predicted by the SVM to be positioned landwards of the coastal vegetation line. The manually digitised line is in red.

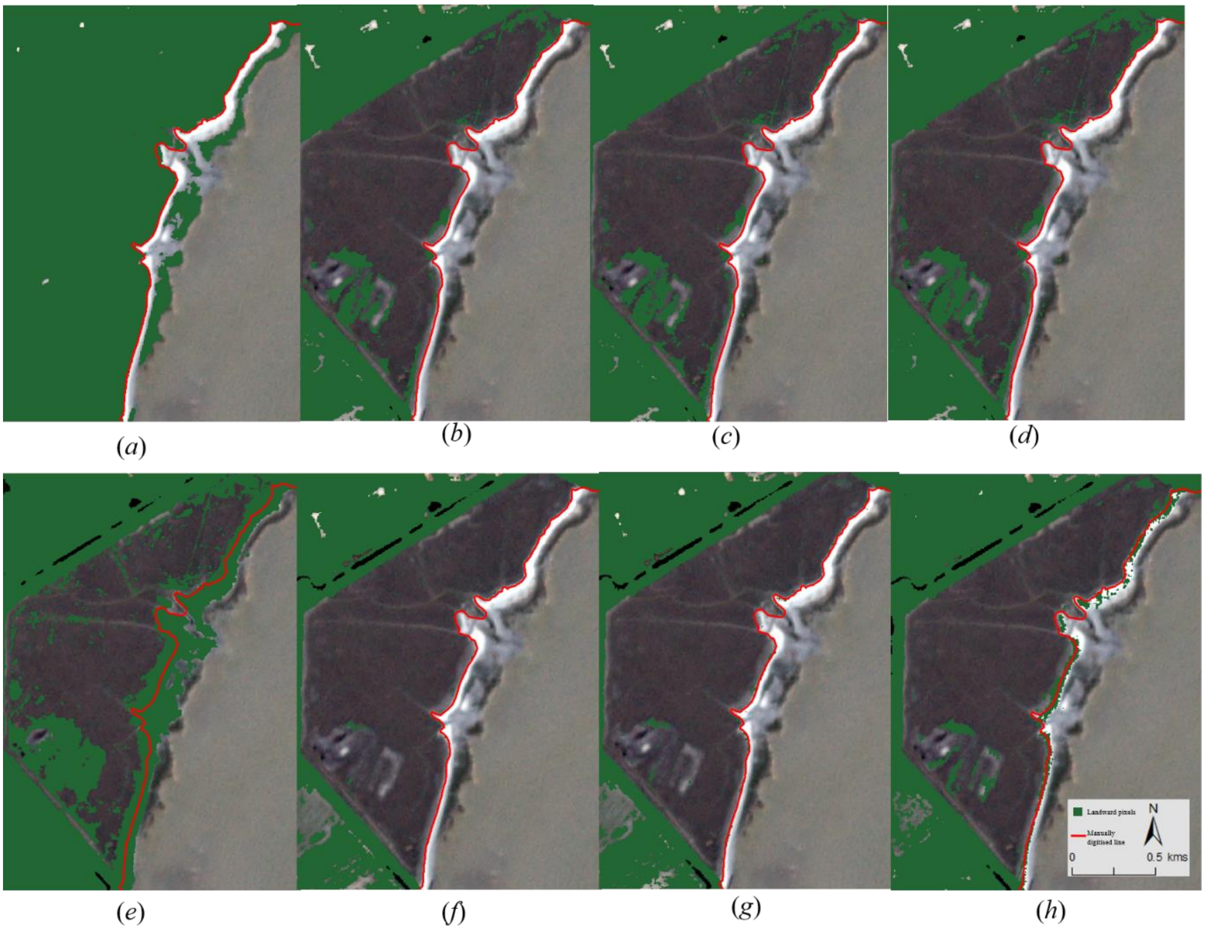


Figure N: Support Vector Machines (SVM) outputs from Hornsey, Essex, UK. SVM output using a linear model and regularisation parameter value, C, of (*a*) 1, (*b*) 2, (*c*) 3 and (*d*) 4. SVM output using a polynomial kernel and a regularisation parameter value, C, of (*e*) 1, (*f*) 2, (*g*) 3 and (*h*) 4. For all images, the pixels in green correspond to pixels predicted by the SVM to be positioned landwards of the coastal vegetation line. The manually digitised line is in red.

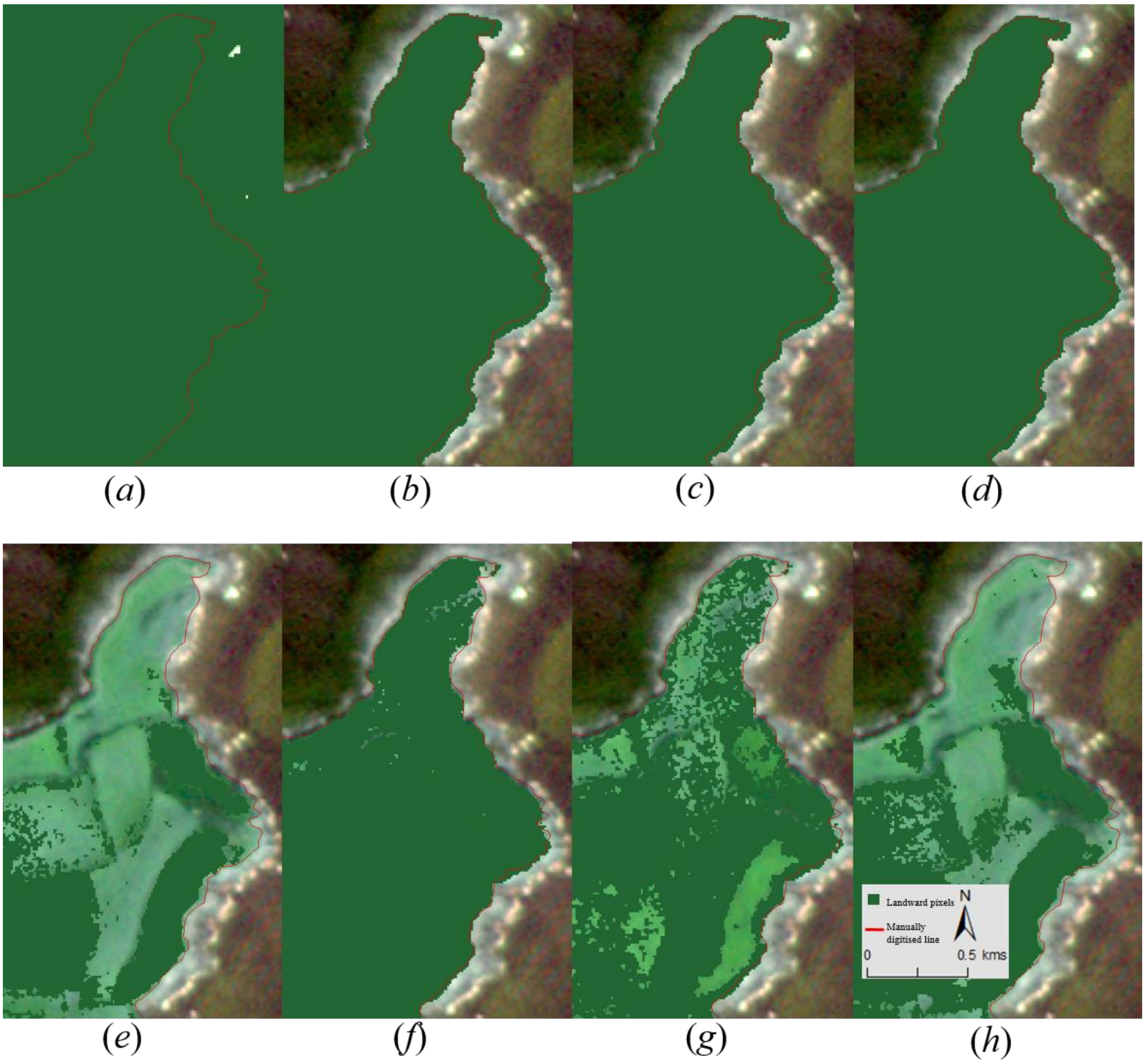


Figure O: Support Vector Machines (SVM) outputs from Porthallow, Cornwall, UK. SVM output using a linear model and regularisation parameter value, C, of (*a*) 1, (*b*) 2, (*c*) 3 and (*d*) 4. SVM output using a polynomial kernel and a regularisation parameter value, C, of (*e*) 1, (*f*) 2, (*g*) 3 and (*h*) 4. For all images, the pixels in green correspond to pixels predicted by the SVM to be positioned landwards of the coastal vegetation line. The manually digitised line is in red.

Supplemental Materials B

Images used during Holistically-Nested edge detection training

Table 1: Locations of convolutional neural network training images. Other columns provide information on dominant shoreline direction, spring and neap tidal ranges, dominant sediment type, geomorphology and climate at each site.

Location	Country	Dominant shoreline direction	Tidal Range (m)	Dominant vegetation type	Geomorphology	Climate
Dunwich-Covehithe	UK (Suffolk)	East	2.5 (spring) 0.5 (neap)	Psammosere dune vegetation/ cliff top grasses and arable crops	Soft cliff &beach dune	Temperate
Winterton	UK (Norfolk)	East	2.5 (spring) 0.5 (neap)	Psammosere dune vegetation	Dune	Temperate
Braunton	UK (Devon)	North West	10 (spring) 3.5 (neap)	Cliff top grasses and Psammosere dune vegetation	Beach dunes and rocky cliff	Temperate
Perranuthnoe	UK (Cornwall)	West	6 (spring), 2 (neap)	Cliff top grasses and Psammosere dune vegetation	Beach dunes and rocky cliff	Temperate
Jacksonville	USA (Florida)	South East	2.5 (spring) 1 (neap)	Scrub (scrub holly/ scrub plum)	Barrier island	Sub-tropical
Ustronie Morskie	Poland	North West	0 (spring) 0 (neap)	Dune vegetation and dense oak and ash woodland	Beach dunes	Temperate

Akwidaa	Ghana	South	2 (spring) 0.5 (neap)	Tropical/ palm forest	Beach dunes	Tropical
Comillas	Spain	North	5 (spring) 1.5 (neap)	Cliff top grasses	Beach dunes and cliffs	Temperate
Bribie Island	Australia	East	2 (spring) 0.5 (neap)	Eucalyptus forest	Barrier island	Sub-tropical
Holderness	UK (East Yorkshire)	East	6 (spring) 1 (neap)	Psammosere dune vegetation/ cliff top grasses	Soft cliff & beach dune	Temperate
Itacare	Brazil	South East	2.5 (spring) 1 (neap)	Tropical broadleaf forest/ palm trees/ retinga vegetation	Beach dune and isolated cliffed headlands	Tropical
Capbreton	France	West	5 (spring) 1.5 (neap)	Psammosere dune vegetation	Dune	Temperate