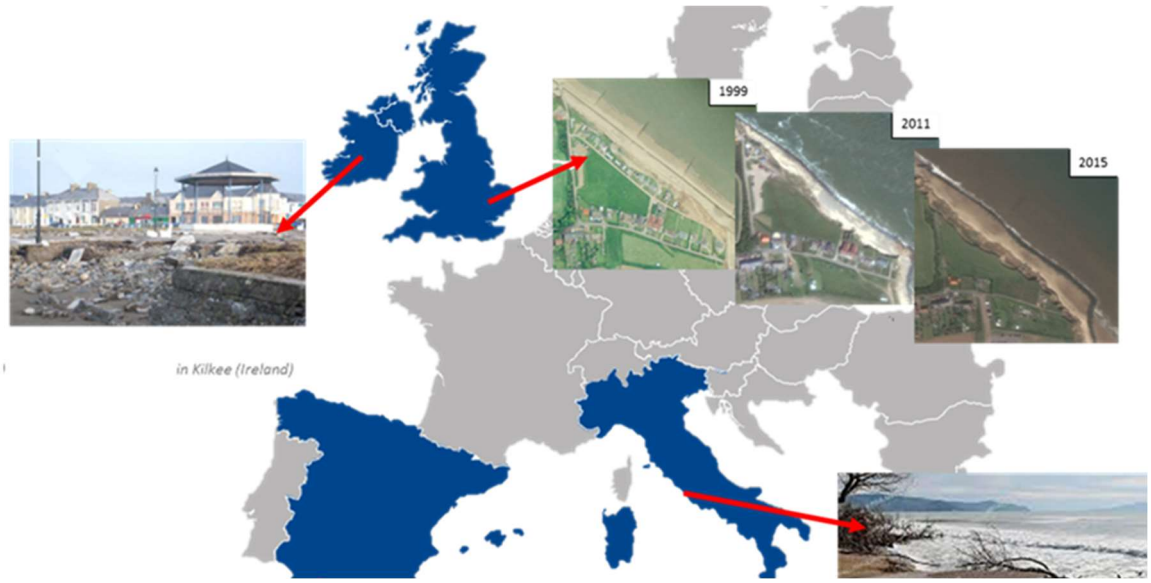
Coastal Erosion from Space



Algorithm Theoretical Baseline Document

Ref: SO-TR-ARG-003-055-009-ATBD-WL

Date: 31/08/2022

Customer: ESA

Contract Ref.:4000126603/19/I-LG

This page is intentionally left blank

Version and Signatures

Version	Date	Modification
Version 1	19/09/2019	–
	30/09/2019	Review and rewriting section1
	03/10/2019	Review
	10/10/2019	Section2 and modification of section1
	20/10/2019	Review
V 1.1	25/11/2019	New adding
	06/12/2019	Review
V 2.0	23/12/2020	Final amendments
V 3.0	10/02/2022	Amendment linked with the processor third version
Verification by	François-Regis Martin-Lauzer	
Authorisation	Craig Jacobs	

Acronyms

CSI: Coastal state indicators

EO: Earth observation

GNDVI: Green normalized difference vegetation index

HR: High resolution

NDVI: Normalized difference vegetation index

NDWI: Normalized difference water index

PSF: Point spread function

ROI: Region of interest

SAR: Synthetic-aperture radar

SDB: Satellite Derived Bathymetry

SNR: Signal to noise ratio

SI: Shoreline indicators

TSM: Total Suspended Matter

VHR: Very high resolution

WL: Waterline

WW: White water

CRS: Coordinate reference system

Applicable and reference documents

Id	Description	Reference
AD-1	Requirement Baseline Document	SO-RP-ARG-003-055-006-RBD_v1.0_20190916
AD-2	End-Users Validation Document	TR_OR_20_018
AD-3	Waterline ATBD V2	SO-TR-ARG-003-055-009-ATBD-WL-VNIR_V2
AD-4	Web Services Specification	SO-RP-ARG-003-055-008-WEB
AD-5	Geolocation ATBD	SO-TR-ARG-003-055-009-ATBD-GL
AD-6	Shoreline Extraction ATBD	SO-TR-ARG-003-055-009-ATBD-SL
AD-7	Technical Specification Document	SO-TR-ARG-003-055-009-TSD

Contents

Version and Signatures ---- **1**

Acronyms ---- **2**

Applicable and reference documents ---- **3**

List of Figures ---- **5**

1 Overview and Background Information ---- **7**

1.1 Product requirement ---- 7

1.2 Quick Review – Feasibility ---- 9

1.3 Product Specifications ---- 16

2 Algorithm Description ---- **18**

2.1 Data Processing outline ---- 18

2.2 Algorithm Input ---- 20

2.3 Theoretical Description of the models in background of the procedure ---- 20

2.4 Algorithm output ---- 28

2.5 Algorithm Performance Estimates ---- 3

3 Conclusion ---- **6**

3.1 Assessment of limitations ---- 6

3.2 Mitigation ---- 7

4 References ---- **8**

List of Figures

Figure **Description** **Page**

Figure 1.1: Diagram of the image processing chain and data flow in Liu & Jezek (2004). ---------- 11

Figure 1.2: Computed line vector relative to surrounding pixel centre points following the Marching Squares algorithm. ---------- 12

Figure 1.3: Noise contained in an otherwise accurate waterline in Start Bay (07/10/2018)---------- 13

Figure 1.4: Waterline detected by the V3 processor in Start Bay (15/01/2020) ---------- 14

Figure 1.5: NDVI white waves respond on a fitted histogram (left) of an RGB image (right)**Erreur ! Signet non défini.**

Figure 1.6: Relation between spatial and temporal scales of morphological features and fluid motions associated with sandy beaches. Source: © Masselink and Kroon ---------- 16

Figure 2.1: Sketch of the computer program ---------- 18

Figure 2.2: Horizontal shift of 1 pixel as a result of co-registration on S2A L2 image of El Prat Barcelona 24/03/2019. Red is non co-registered and purple from co-registered data.---------- 19

Figure 2.3: Spectral reflectance curves for vegetation, soil and water (Lillesand, Kiefer, & Chipman, 2004)21

Figure 2.4: Definition of terms related to tides ---------- 22

Figure 2.5: Zoom of an NDVI image over a coast area in Barcelona (left) and cross shore profile corresponding to the pins' values (right) ---------- 24

Figure 2.6: Spectral profile of sea (blue) and land –mix land cover (green) of Barcelona ---------- 25

Figure 2.7: Waterline performance differences for different spectral indices: NDVI (red), GNDVI (green), BNDVI (cyan). Barcelona (tile: 31TDF), 02/07/2019. ---------- 26

Figure 2.8: Illustration of error and uncertainty. Error in measurement (purple), uncertainty (green), (a) conventional random error, (b) systematic component error, cannot be characterized with a single value (Povey and Grainger, 2015) ---------- 27

Figure 2.9: Internal quality control, noise correctly identified by lower scores (red and orange). ---------- 2

Figure 2.10: External quality control – darker colours representing noise, and the bright line around the coast showing the density of waterlines where we would expect to detect them.---------- 2

Figure 2.11: V3 waterline (cyan) against 10m buffer around V2 waterline (red) for a section of (a) Barcelona and (b) rugged coastline in Start Bay.---------- 4

Figure 2.12: V3 waterline (cyan) and V2 waterline (red) for a port in Barcelona ---------- 5

Figure 2.13: Clouds, white water and passing airplane causing discontinuities and false edges in BNDVI-based (a) V2 waterline processor compared to (b) BNDVI-based V3 waterline processor, Barcelona, 21/08/2019 S2 tile T31TDF ---------- 5

1 Overview and Background Information

1.1 Product requirement

The overall objective of the Coastal Erosion from Space project is to retrieve Coastal State Indicators (CSI) which describe the dynamic-state and evolutionary trends of coastal systems, from Shoreline Indicators (SI), gauges, pointers or markers that are used as proxies to represent the shore (either visible discernible features, or tidal datum-based indicators) to get isobaths and isohypses (contour lines) (Ref: SO-RP-ARG-003-055-006-RBD_v1.0_20190916, AD-1).

Coast, coastline, shore and shorelines are often confused, although the focus is on the same geographical features:

- a shore or a shoreline is the fringe of land at the edge of a body of water (the 'line' is quite thick on a large-scale map, e.g. 1:5,000, but very thin on a small-scale map, e.g. 1:1,000,000. Thus, a precise line that can be called a shoreline cannot be determined if it does not refer to a representation scale or spatial frequency cut),
- a coast, also called coastline or seashore, is a shore which borders the sea; however, coast often refers to an area far wider than the shore, often stretching miles into the hinterland.

1.1.1 *Information content quality and value*

1D EO products, shorelines from waterlines, are designed by identifying points in the digital images at which the image characteristics change sharply when coming landward from the sea. Those shorelines are needed to map and study the dynamics of the land/water interface whether visible or not to the human eye.

Land/water delimitation is generally equated with the maritime boundary; however, all land/water edges are not necessarily maritime. We have the example of lacs, riverbanks or offshore sand bars that may be underwater most of the time. We here consider as waterline, each land/water interface connected to the sea and inside our "coastal area", i.e. buffer around the coast which delimits our area of interest for the project. We thus consider riverbanks near the coast if they are open to the sea, and offshore sand bars when they are above water.

1.1.2 *End user feedback*

Version 2 of the waterline processor was tested and validated by the end-users, who then provided feedback (Ref: TR_OR_20_018, AD-2). Overall, previous versions of the waterline processor generated good results with accurate representative waterlines to the satisfaction of the end users. The main points of feedback to address with a new processor were:

- a “jigsaw” pattern found in the waterline due to the line traveling along pixel edges,
- discontinuous waterline segments,
- fictitious waterline from sediment, foam spots, and features submerged in clear water
- small lakes and channels near the true coastline being detected.

Limitations of the previous versions were also identified during development and testing. For more information on these please refer to the relevant ATBDs (Ref: SO-TR-ARG-003-055-009-ATBD-WL-VNIR_V2, AD-3).

The version 3 of the waterline processor follows the same general approach as the previous versions, applying different methods for the kernel centre location and derivation of the waterline vector from the binary land and sea mask. It is expected that these alterations will address the main points of concern highlighted in the end user feedback.

1.1.3 *Product order & delivery services*

The 1D proxy-based shoreline indicator (waterline) product is based on 25-years of EO data. We use different spectral properties of optical data sets as reported for studies exploiting data at different resolutions (namely LANDSAT 5 and 8 at 30m and SENTINEL-2 at 10m, and also VHR (including SPOT, IKONOS, PLEIADES and Worldview at 5m or less).

Every waterline is delivered to a specific naming convention format. The naming convention includes identification information about each waterline including; a coastal erosion identifier (CE), a start date, type (eg . VNIR or SAR waterline), category, level, coordinates, sensor qualifier, production date and the extension (.tif). All codes are separated by an underscore.

The following template is used:

CE_<startdatetime>_<type>_<category>_<level>_<bbox>_<qualifier>_{<enddate>}_<procdate>.<extension>

An example waterline:

e.g. CE_20200221110959_WL_OB_L2_360021N-63636S-370148N-53339S_S2_20220126.geojson.

For full information regarding the naming convention, please refer to the Web Services Specification (Ref: SO-RP-ARG-003-055-008-WEB, AD-4).

The waterline product includes the delivery of:

- decadal geomorphological changes over a 25-year period based on waterlines (= proxy-based shoreline indicators).
- a feasibility study with regards to monitoring the shoreline with the Sentinel HR constellation, or other public satellites, and only with commercial VHR satellites for large scale observations as appropriate.

Delivery as a GeoJSON file for GIS software.

1.2 Quick Review – Feasibility

There have been a number of approaches to proposing methods for defining proxy waterlines based on multispectral EO products. Various image analysis methods have been applied, for instance change detection and different classification procedures. Several of them relied on band ratios, such as the NDVI, as input parameters for classification to aid the definition of boundaries between different categories of land and between the land and sea as well[1].

1.2.1 *Satellite sensors and mission*

Please refer to the Geolocation ATBD (Ref: SO-TR-ARG-003-055-009-ATBD-GL, AD-5).

[1] Muttitanon, W. & author, N. K. T. C. Land use/land cover changes in the coastal zone of Ban Don Bay, Thailand using Landsat 5 TM data. *International Journal of Remote Sensing* **26**, 2311–2323 (2005).

1.2.2 *Models specification*

Land cover identification is usually done by classification[2,3], however, due to the spectral characteristics of water, other methods are more effective for the identification of the interface between land and water. At first, waterline detection was done manually by operators on aerial photography, now remote sensing allows for the detection of features over larger areas at the same time, and multi-spectral and multi-temporal analysis improves the detection process. With aerial photography, waterline was identified by eye using the available spectra (black and white or colour). The waterline was depicted by a change in the grey tone, and a correction was applied to remove tidal effects[4]. The generation of commercial high-resolution satellite imagery allows an improvement in shoreline indicators mapping, spatial resolution is comparable to that of aerial photographs for a lower cost, and the short revisit time allows stereo mapping. The instantaneous waterline can be extracted from a single band image, due to water's strong spectral absorption. A histogram thresholding process carried out on one infra-red band will separate water bodies (low values) and land (high values). However, finding the exact value at which land begins has remained an issue.

On the other hand, a threshold on a band ratio histogram instead of on a single band histogram is more marked, water and land are directly separable[5]. Usually, brightness is used, but also spectral indices such as NDVI or NDWI. Another approach for shoreline extraction is to realize a segmentation to partition the input image into homogeneous regions with a locally adaptive threshold. Then a differentiation between coast edge and other[6] regions are carried out. The variety of mathematical methods based on gradient operators falls under the name "edge detection". The result of applying an edge detector leads to a set of +/- connected curves that indicate the boundaries of the sea.

---

[2] Keuchel, J., Naumann, S., Heiler, M. & Siegmund, A. Automatic land cover analysis for Tenerife by supervised classification using remotely sensed data. Remote Sensing of Environment 86, 530–541 (2003).

[3] Mohammady, M., Moradi, H. R., Zeinivand, H. & Temme, A. J. A. M. A comparison of supervised, unsupervised and synthetic land use classification methods in the north of Iran. *Int. J. Environ. Sci. Technol.* **12**, 1515–1526 (2015).

[4] Stafford, D. B. Air Photo Survey of Coastal Erosion. *Photogrammetric Engineering & Remote Sensing* 11 (1971)

[5] Alesheikh, A. A., Ghorbanali, A. & Nouri, N. Coastline change detection using remote sensing. *Int. J. Environ. Sci. Technol.* **4**, 61–66 (2007).

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

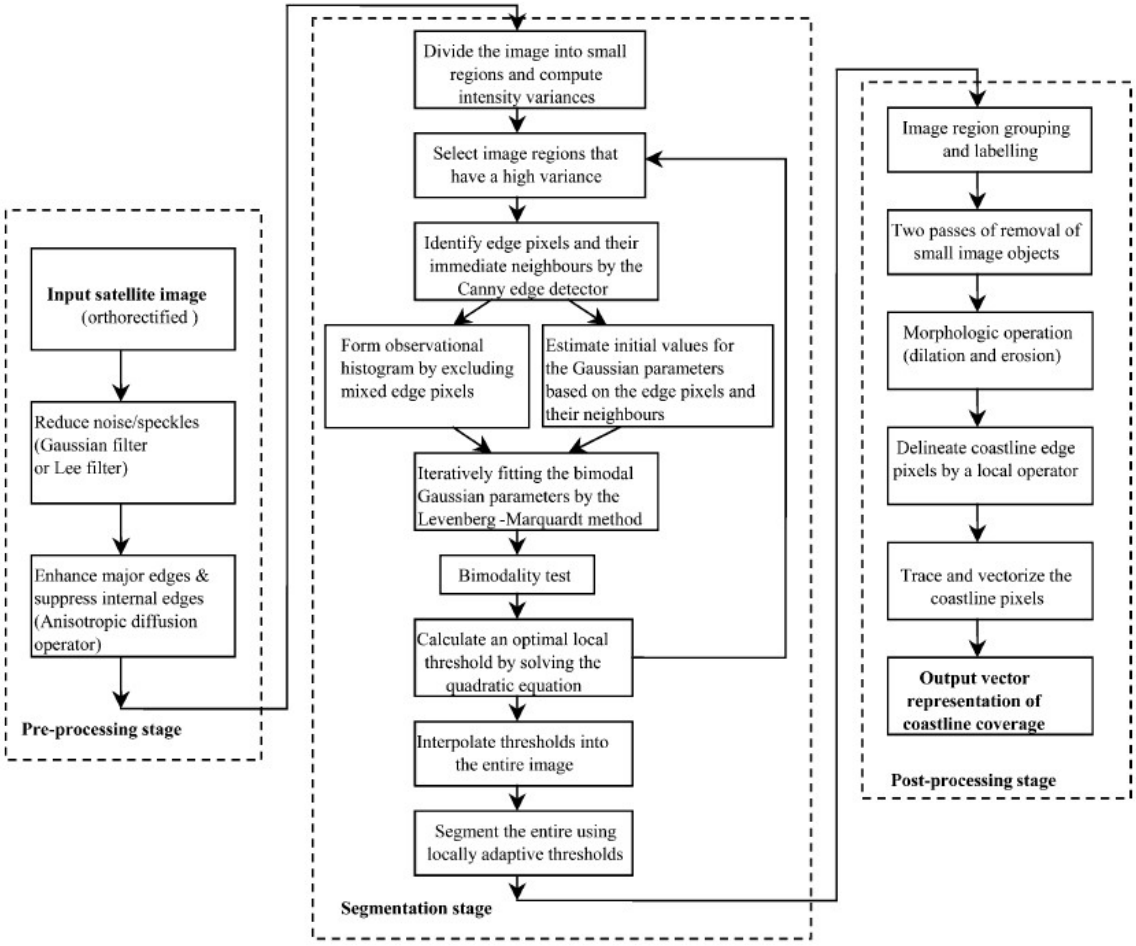


Figure 1.1: Diagram of the image processing chain and data flow in Liu & Jezek (2004).

Various vectorisation approaches are detailed in different studies. To transform raster features into a vector file the gdal contour[7] method implements a Marching Squares algorithm to draw a line through pixels using the neighbouring pixels value to estimate the position of the line as shown in **Figure 1.2**. These values are determined by the user-defined elevation interval, which is used to classify pixels into equally sized intervals based on the pixel value.

[7]Warmerdam, F. & Rouault, E. a. o. gdal_contour. *GDAL documentation,* [Online] Available at: https://gdal.org/programs/gdal_contour.html (2022) [Accessed 21 February 2022].

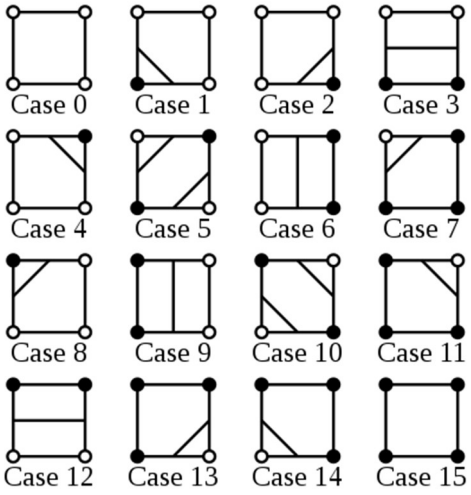


Figure 1.2: Computed line vector relative to surrounding pixel centre points following the Marching Squares algorithm.

1.2.3 *Auxiliary data*

To make the process more efficient the computation of EO-based waterlines benefits from the addition of an approximate global coastline vector[8] – e.g. country or coastline shapefiles available for the different areas of interest. Additionally, the area of interest must be specified by a shapefile input. The coordinate reference system (CRS) of the global coastline must be the World Geodetic System 1984 (WGS 84, EPSG: 4326), which is the most appropriate geodetic datum for data with global coverage. EO products and the auxiliary area of interest shapefile can have any CRS, and need not match. The generated waterline will be in the CRS of the EO product input.

1.2.4 *Currently known issues*

Obtaining a continuous waterline from EOs is not always possible, even with moderate complexity, i.e. taken in calm weather, because of shoals, cusps, etc. Although band ratios, such as NDVI, have been useful classification tools, these discontinuity issues with extracted edges hampered by fragmentation have surfaced in the context of waterline delineation in relation to land use/land cover mapping[9] and coastline change

[8] National Geospatial-Intelligence Agency, 2022. Digital Nautical Chart (DNC®). [Online] Available at: https://dnc.nga.mil/dncp/home.php (2022) [Accessed 25 February 2022].

[9] Muttitanon, W. & Tripathi, N. K. Land use/land cover changes in the coastal zone of Ban Don Bay, Thailand using Landsat 5 TM data, *International Journal of Remote Sensing*, 26:11, 2311-2323 (2005).

detection[10,11] for instance. The V3 waterline sees a drastic improvement to the continuity of the waterline, for example, a V2 waterline derived from an image in Start Bay taken on 27/09/2018 contained 284 line segments with a mean line length of 479m, whereas the corresponding V3 waterline covered the same distance with only 131 segments and a mean line length of 1122m. However, with the improvement in continuity, a new issue of the V3 waterline processor is observed where noise is contained in a long, otherwise accurate waterline as shown in Figure 1.3.

Figure 1.3: Noise contained in an otherwise accurate waterline in Start Bay (07/10/2018)

One of the reliable interesting uses of EO is the drawing of the instantaneous interface between the water and +/- dry materials of the land, so-called “waterline” (WL), with the main difficulties being:

- different features being detected based on the time of image, such as shoals and sand bars that may be visible depending on the tide height.
- waterline may be detected as the interface between white water and calmer sea, rather than white water and land (Figure 1.4: Waterline detected by the V3 processor in Start Bay (15/01/2020)Figure 1.4).

[10] Maglione, P. et al. Coastline extraction using high resolution WorldView-2 satellite imagery, European Journal of Remote Sensing, 47:1, 685-699 (2014).

[11] Kumer Ghosh, M. et al. Monitoring the coastline change of Hatiya Island in Bangladesh using remote sensing techniques, ISPRS Journal of Photogrammetry and Remote Sensing, 137–144 (2015).

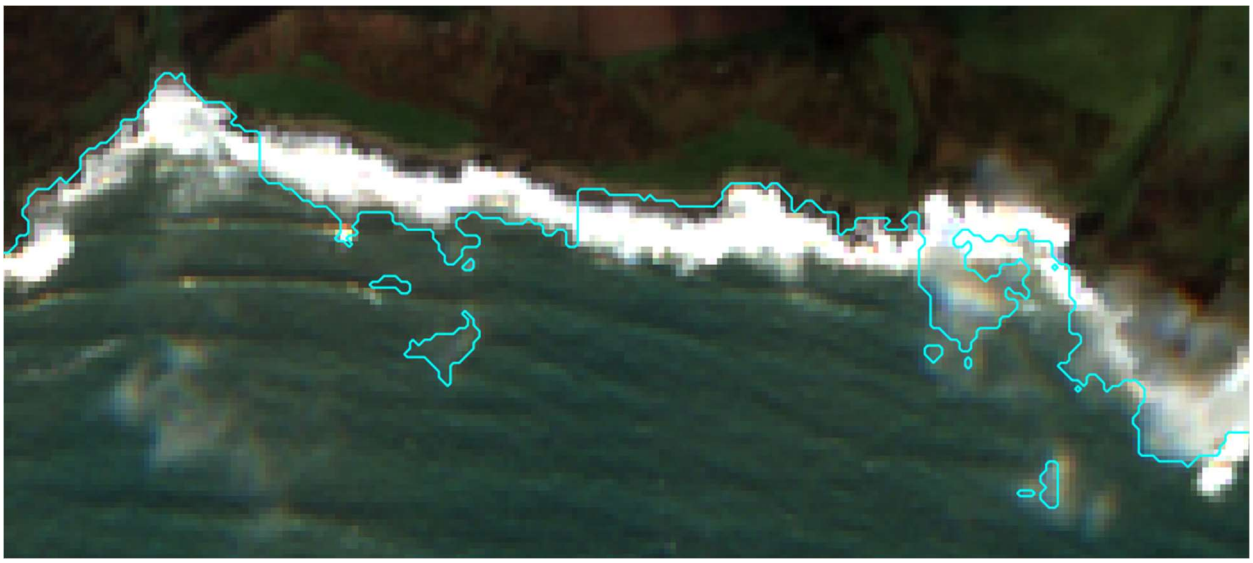


Figure 1.4: Waterline detected by the V3 processor in Start Bay (15/01/2020)

The spectral response from white water can be closer to that of land than calm water (**Erreur ! Source du renvoi introuvable.Erreur ! Source du renvoi introuvable.**) which leads to the observed issue of white water clustered with land and the waterline being generated between white water and calm sea.

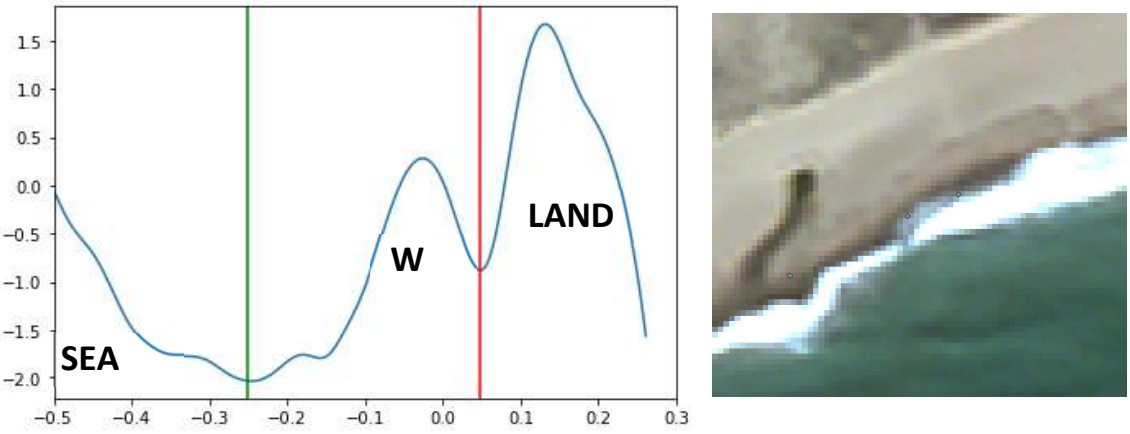


Figure 1.5: NDVI white waves respond on a fitted histogram (left) of an RGB image (right)

A waterline derived from EO is not only prone to observation errors but cannot be considered as an isobath or a contour line (isohypse), due to a. the waves reaching different height on the shore at different points (wave set-up of the surf zone, wave run-up of the swash zone), and b. the tide height being different along the shore because of shore morphology; as such it is critical to have an error/uncertainty budget.

1.2.5 *Potential solutions to be developed*

There are two possible options for white water detection being considered:

- White water values in RGB and NIR are commonly greater than three standard deviations above the tile's/scene's mean values for each band, therefore placing them in the brightest 0.3% of the image. These bright pixel locations are mapped, and an adaptation to the processor can be introduced if the percentage of bright pixels is too high. Above this value we would expect to see a peak between those of land and sea and thus we need to provide an additional constraint that we only take the trough value on the landward side of the white-water peak. We could then add in confidence qualifiers within metadata to say that due to the presence of white water, we have a lower confidence in the accuracy of the line.
- The brightest 0.3% of pixels are removed, and these regions are then filled in using values from the surrounding pixels. A threshold is then derived from the region, which has now had the brightest white water removed, using the standard method. As long as the bright regions are completely surrounded by either land or sea, then they will be filled in correctly.

To account for the variability of waterlines at a single site at different times of the day and year caused by morphology and tidal variability (Figure 1.6), auxiliary tidal and elevation data will be integrated with the waterlines taking into account the date and time of the image from which they were derived to produce shorelines. Information on this can be found in the Shoreline Extraction ATBD (Ref: SO-TR-ARG-003-055-009-ATBD-SL, AD-6)

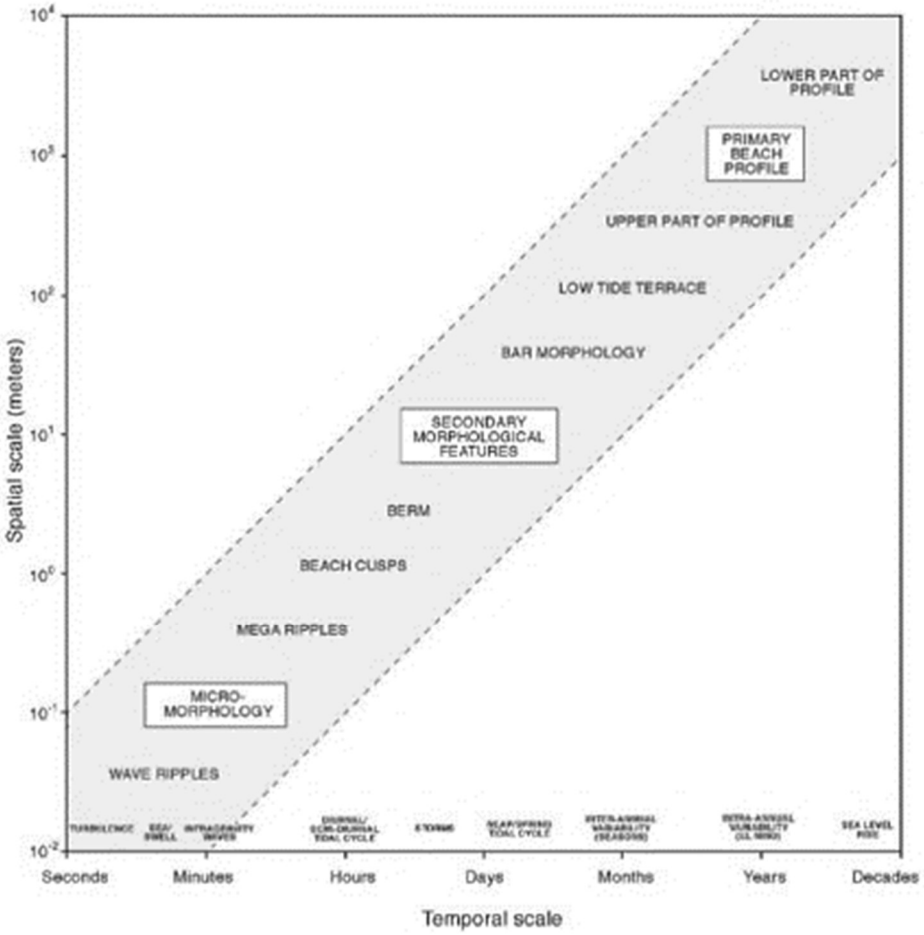


Figure 1.6: Relation between spatial and temporal scales of morphological features and fluid motions associated with sandy beaches. Source: © Masselink and Kroon[12]

1.3 Product Specifications

The shoreline indicator coastline is produced by the addition of in-situ information such as sea state, meteorological forecast, tide information to a waterline which is the instantaneous boundary between water and land. Waterlines will be computed from EO optical products (HR and VHR) using a locally adaptive thresholding method on spectral indices: NDVI, BNDVI, GNDVI, etc. according to the type of land cover (building on the established knowledge of the operator and the local specificities of the study locations).

The thresholding method works on the assumption of a bimodal distribution of pixel values representing land and sea to find the value in between peaks to act as the threshold value. However, this may not always the case, for example, where the presence of white water introduces a third peak resulting in the algorithm

[12] COASTAL ZONES AND ESTUARIES – Morphology and Morphodynamics of Sandy Beaches - G. Masselink, A. Kroon

choosing an inaccurate threshold. Therefore, a threshold range and default threshold value is specified by the user for use where the pixel value histogram is not bimodal.

The waterlines will be extracted annually or seasonally according to erosion rate to develop regular monitoring tools; and around storm-events to improve short-term response and emergency works. The waterlines will be vectorized and will be available on a geoportal as ready-to-use products. Every waterline is produced with an associated internal quality control (QC) score within the shapefile product. The internal QC analyses the geometric characteristics of the waterline in segments and assigns a score between 0 and 100 (0 = low confidence, 100 = high confidence), the higher the score, the greater the confidence that the line is not erroneous. This is discussed in more detail in section 2.

2 Algorithm Description

2.1 Data Processing outline

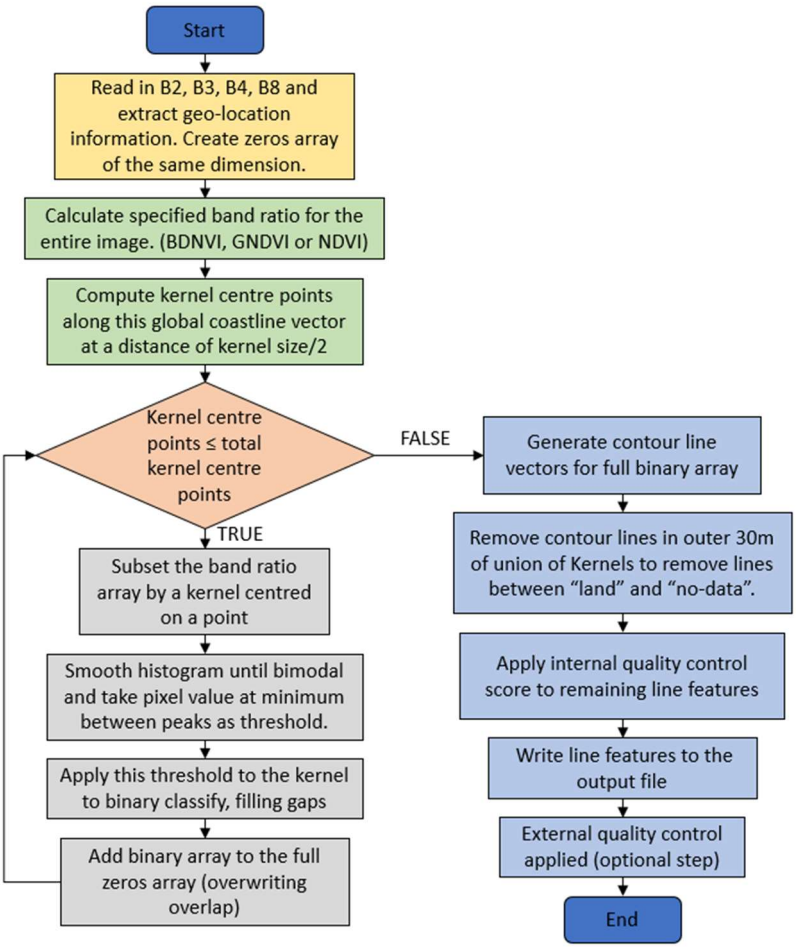


Figure 2.1: Sketch of the computer program

The waterline processor is designed to process data from different sensors, and a standardisation step is needed to fit data from different sensors to the right format. The user chooses one of the NDVI, GNDVI, BDNVI or other band ratios to apply to the image. An auxiliary file containing the global coastline vector files is used as a "guide" along which a small kernel window slides to produce the waterline using an adaptive threshold. To calculate this adaptive threshold, the histogram of index pixels inside a kernel is smoothed to a bimodal distribution, with the minimum value between the peaks taken as the threshold to differentiate between land and water. If this threshold is outside the specified threshold range, then the user-defined default threshold value is used instead to mitigate against false edges. Each individual kernel is added into a full area of interest

binary array of land and sea, from which contouring methods extract the produced waterline and write the waterline in vector format. Finally an internal QC score is calculated and written into the vector file metadata.

2.1.1 *Pre-requisite*

As we realise a comparative study over decades, all images need to be co-registered from a master image. An example of co-registration can be seen below (Figure 2.2), where the red line represents the waterline computed from a non-co-registered image and the purple line from a co-registered image overlaid on the latter. As a result, a horizontal shift of one pixel can be observed, implying a magnitude of 10m adjustment in the final waterline product due to the pre-processing. For further details please refer to the Geolocation ATBD (Ref: SO-TR-ARG-003-055-009-ATBD-GL, AD-5).

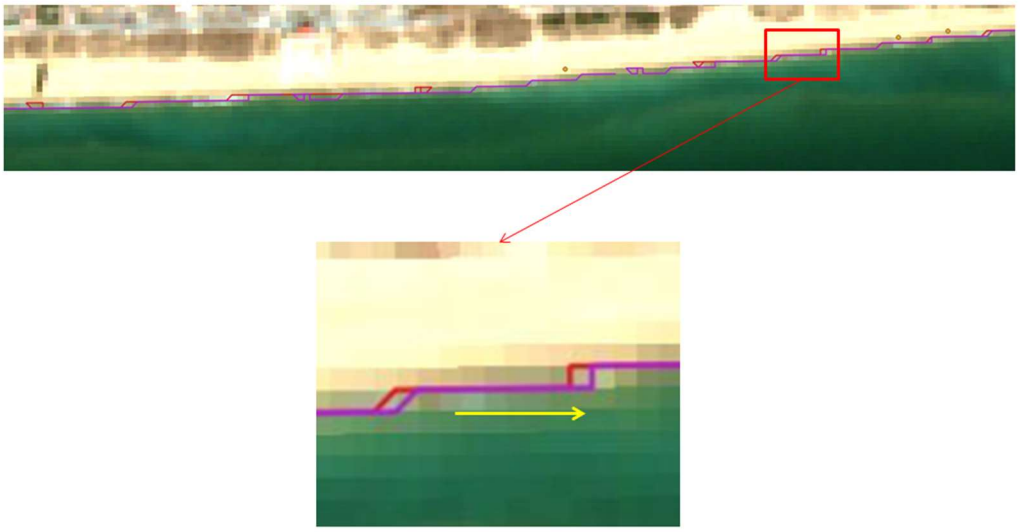


Figure 2.2: Horizontal shift of 1 pixel as a result of co-registration on S2A L2 image of El Prat Barcelona 24/03/2019. Red is non co-registered and purple from co-registered data.

The waterline processes use spectral indices, NDVI, BNDVI and GNDVI to extract waterlines by using different parts of the EM spectrum which is captured by the sensor to create spectral indices. A series of tests are performed to see how the waterline processor performs based on the results of the binary classification of the region of interest (ROI) into land and water. This binary classification is based on one of the band ratios: NDVI, BNDVI or GNDVI. The best performing spectral indices, in term of the position of waterline, is selected. This is analysed in several ways, such as with a visual check: this involves looking at the performance and considering characteristics such as cloud interruption, rough-white waters, and tide levels.

It is important to ensure that the processor extract the land/ sea boundary and no other features like wet sand and vegetation.

Metric comparison can also be used to analyse the performance of spectral indices by calculating the percentage of areal coverage of each performing range.

2.2 Algorithm Input

- Path to directory containing co-registered optical image(s) (HR or VHR)
- Output directory path
- Band ratio (NDVI, GNDVI or BNDVI)
- Kernel size (metres)
- Minimum and maximum threshold values
- Default threshold value
- Location name (for output metadata)
- Global coastline vector files from auxiliary data
- Polygon shapefile identifying the region of interest

2.3 Theoretical Description of the models in background of the procedure

2.3.1 *Physical Description*

If the boundary between land and water is quite easy to detect by eye, this line may also be detected using satellite imagery due to the spectral characteristics of water bodies.

Spectral profiles for bare soil, green vegetation and clear water are available on Figure 2.3. Water bodies are characterized by small reflectance in visible spectrum and a (quasi-) total absorption in higher wavelengths.

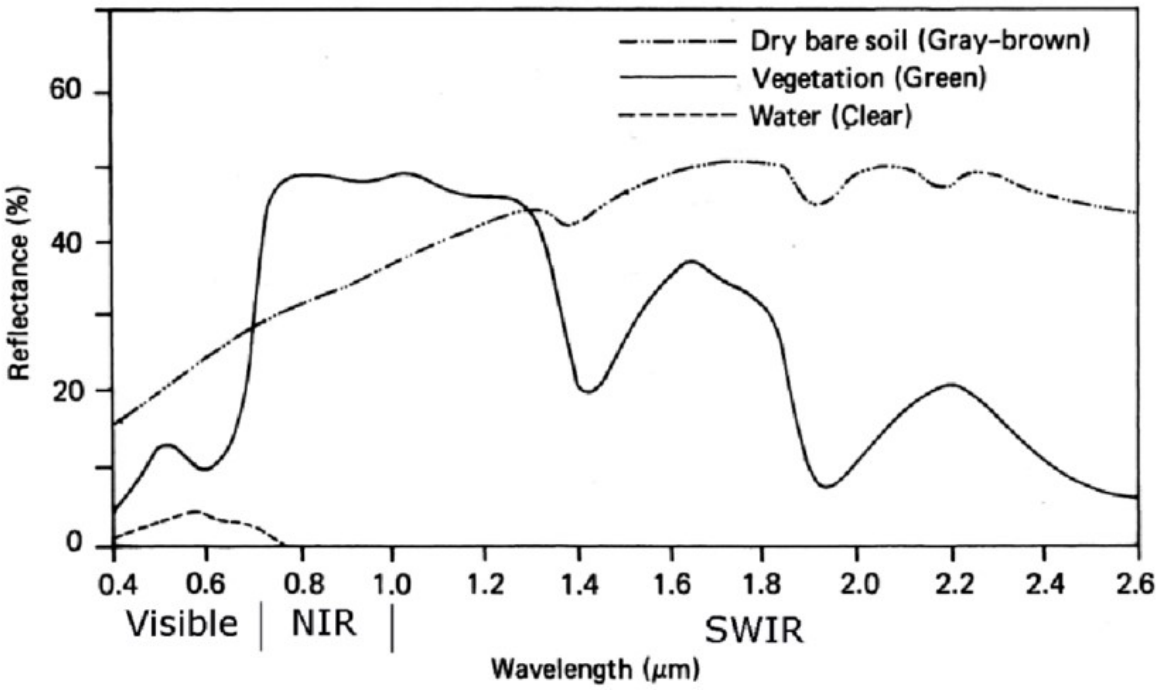


Figure 2.3: Spectral reflectance curves for vegetation, soil and water (Lillesand, Kiefer, & Chipman, 2004)

NIR band, SWIR band and even band ratios using those spectral differences should allow for an accurate waterline extraction.

2.3.2 *Mathematical Description and calculation procedures*

A waterline is the border between water and land at a given time t in a given area– an instantaneous border:

$$WL(t) = \left\{\vec{X}_{\mathrm{WL}}(t,s) | s \in [s_0(t), s_1(t)]\right\}$$

where s is a curvilinear coordinate, $\vec{X} = \left(\vec{X}_h, z\right) | \vec{X}_h = (x, y)$,

z can't be derived from an image or EO (except if we were to perform stereo plotting with a couple of EOs with proper accuracy), contrary to $\vec{X}_h$;

as such $WL^{\mathrm{EO}}(t) = \left\{\vec{X}^{\mathrm{EO}}_{\mathrm{WL,h}}(t,s) | s \in [s_0(t), s_1(t)]\right\}$ and $z_{WL}(t,s)$ is an external/ independent parameter which will be noted $z^{ext}_{WL}(t,s)$; $z^{\mathrm{ext}}_{\mathrm{WL}}(t,s) = z^{\mathrm{ext}}_{\mathrm{WL}}\left(\vec{X}^{\mathrm{EO}}_{\mathrm{WL,h}}(t,s)\right)$ is not a constant, as the altitude of the sea level changes from one place to the other because of the astronomical tides, the prevailing meteorology (atmospheric pressure), and the consequences of meteorological events at the coast (wind waves & swells, that break when reaching the shore).

However, the wave-filtered sea-surface away from the shore is a relative reference, its altitude to the geoid or a reference ellipsoid or reference points ashore being known as astronomical tides are deterministic $z^{\text{ext}}_{0,\text{ocean}}(\hat{s}) + z^{\text{ext}}_{\text{ocean\_as-tide}}\left(t, \widehat{\vec{X}_h} \,|scale\right)$; where ∴ represents spatial scales of the astronomical tides, i.e. a few 1000s km, but they need be modulated by the ocean border geomorphological scales $\delta z^{\text{ext}}_{\text{costal_as-tide}}\left(t, \widehat{\vec{X}_h} \,|scale\ast\right)$such as those of the English Channel, leading to

$$z^{\text{ext}}_{0,\text{ocean}}(\hat{s}) + z^{\text{ext}}_{\text{ocean\_tide}}\left(t, \widehat{\vec{X}_h} \,|scale\right) + \delta z^{\text{ext}}_{\text{costal\_tide}}\left(t, \widehat{\vec{X}_h} \,|scale\ast\right)$$

as one expects an EO tile to be smaller than the length scale of the astronomical tide components, it becomes independent from the position of the waterline: $z^{\text{ext}}_{0,\text{ocean}} + z^{\text{ext}}_{\text{ocean\_tide}}(t) + \delta z^{\text{ext}}_{\text{costal\_tide}}(t)$.

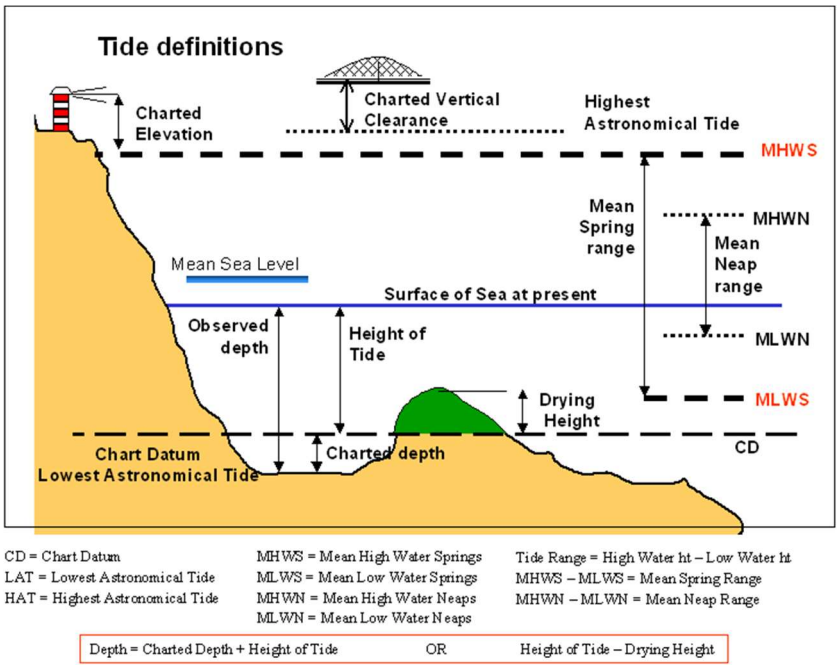


Figure 2.4: Definition of terms related to tides

Nevertheless, meteorological phenomena occurring at different scales provide additional considerations:

- atmospheric ones at scales of O(100 km), being larger than an EO tile, i.e.$\delta z^{\text{ext}}_{\text{atm}}(t)$, incl. wind-driven currents that may pile up water on a coast;
- ocean waves at scales from a few cm to hundreds of km, i.e. $\delta z^{\text{ext}}_{\text{waves}}(t)$, this term having 3 components:

- the effect of Stokes current induced by waves, i.e.$\delta z^{\text{ext}}_{\text{waves,Stokes}}(s,t)$ which is considered as small enough to be discarded;
- the wave set-up or local elevation in the mean water level on the foreshore caused by the reduction in wave height through the surf-zone, i.e. $\delta z^{\text{ext}}_{\text{waves,set-up}}(s,t)$;
- the wave swash or alternate rise and fall of the waterline on the shore slope due the propagation and reflection of waves, i.e. $\delta z^{\text{ext}}_{\text{waves,swash}}(s,t)$.[13]

These result in the following: $z^{ext}_{WL}(t,s) = z^{\text{ext}}_{0,\text{ocean}} + z^{\text{ext}}_{\text{ocean\_tide}}(t|10\text{mn}) + \delta z^{\text{ext}}_{\text{costal\_tide}}(t|\text{few-mn}) + \delta z^{\text{ext}}_{\text{atm}}(t|1\text{h}) + \delta z^{\text{ext}}_{\text{waves,set-up}}(s|10\text{m}, t|\text{few-mn}) + \delta z^{\text{ext}}_{\text{waves,swash}}(s|1\text{m}, t|\text{few-s})$

Considering only the variations:

$$\delta z^{ext}_{WL}(t,s) = z^{\text{ext}}_{\text{ocean\_tide}}(t|10\text{mn}) + \delta z^{\text{ext}}_{\text{costal\_tide}}(t|\text{few-mn}) + \delta z^{\text{ext}}_{\text{atm}}(t|1\text{h}) + \delta z^{\text{ext}}_{\text{waves,set-up}}(s|10\text{m}, t|\text{few-mn}) + \delta z^{\text{ext}}_{\text{waves,swash}}(s|1\text{m}, t|\text{few-s})$$

a waterline $WL^{\text{EO}}(t)$ is defined by $WL^{\text{EO}}(t) = \left\{\left(\vec{X}^{\text{EO}}_{\text{WL,h}}, \delta z^{ext}_{WL}\right)(t,s) | s \in [s_0(t), s_1(t)]\right\}$.

Translating this mathematical concept into the result of a computational process led to the development of the waterline processor algorithm as described in sections 2.1, 2.2 and 2.3.1. The below highlights mathematical procedures concerning the band ratio calculations involved with the algorithm.

Figure 2.5 focus on a coastal area in Barcelona, an NDVI image was computed and pixel values were extracted across the shore to realise a profile. We observe, for sea areas, negative values whereas for land areas values are positive. However, setting a threshold at zero is not enough to extract the sea/land boundary.

[13] Summing up the last two terms results in wave run-up (the maximum level the waves reach up on the beach relative to the still water level)

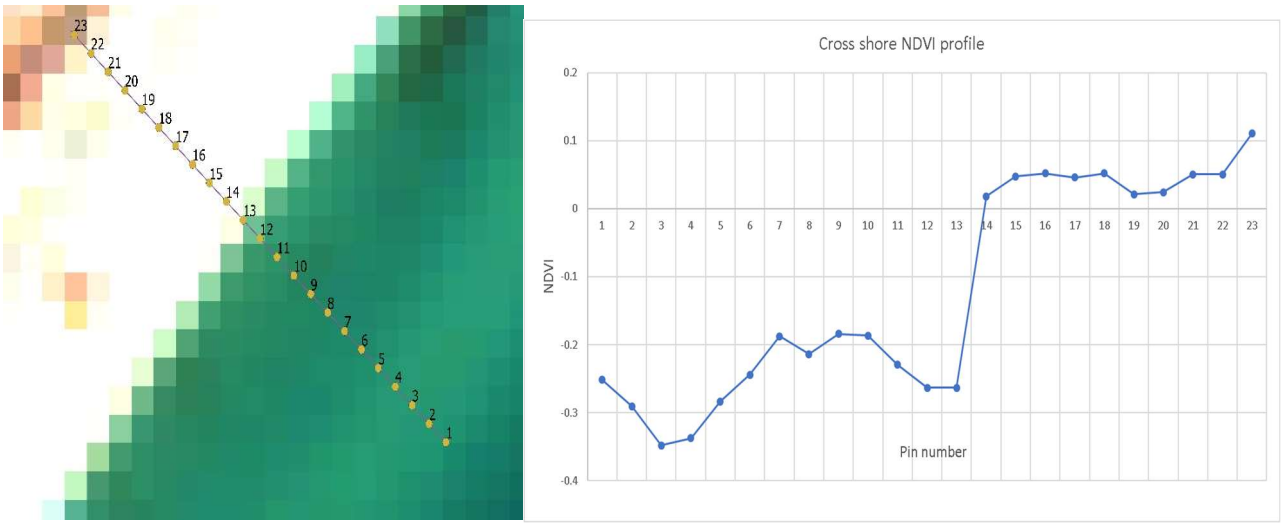


Figure 2.5: Zoom of an NDVI image over a coast area in Barcelona (left) and cross shore profile corresponding to the pins' values (right)

In the RGB image we can see that pin 13 is on the very edge of the water, whereas pin 14 is inland. In the cross shore profile on the right, we can see the change between pin 13 and 14, going from negative to positive values respectively.

2.3.3 *Acceptance of the Models*

Remote-sensing data allow water bodies detection over large areas at a given time, studies have shown that a single band classification using the infra-red band is giving a better representation of water bodies than visible bands. However, some confusions are observed for urban areas and hill shadows. With a multispectral classification, e.g. using band ratios, those confusions can be avoided[14]. Figure 2.6 shows that the separation between build land areas and water bodies may be problematic while using a single band.

[14] Paul Shane Frazier & Kenneth J. Page. Water body detection and delineation with Landsat TM data. *Photogrammetric Engineering & Remote Sensing* vol. 66 1461–1467 (2000).

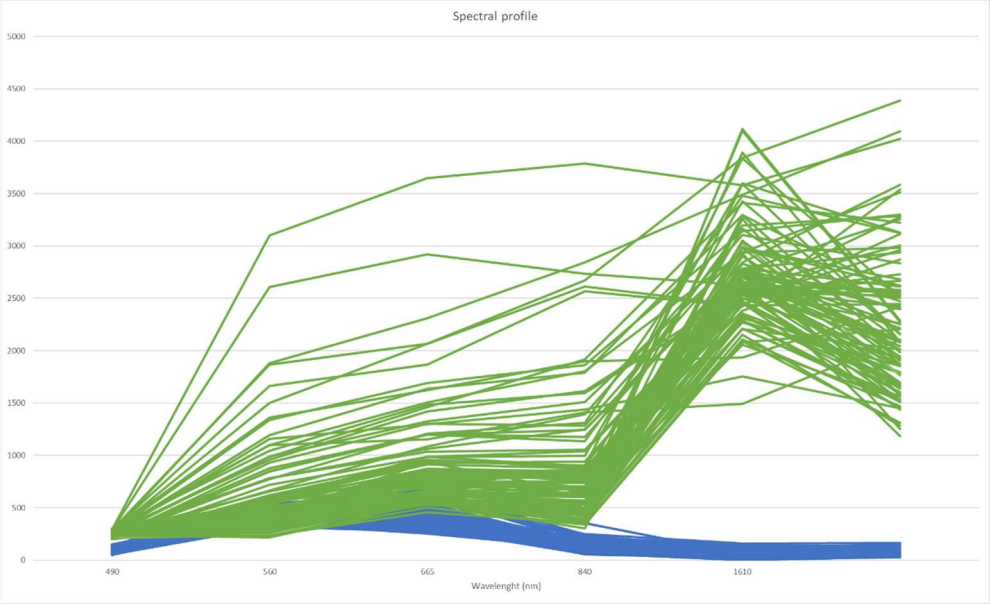


Figure 2.6: Spectral profile of sea (blue) and land –mix land cover (green) of Barcelona

A threshold value based on a single band may produce erroneous waterlines due to the effect of sediment load, the use of band ratios allows to increase the spectral distance between land and sea. A threshold on an NDVI image produces a more accurate line than a threshold on a single band image – NIR or SWIR[15]. Depending on the coast morphology and land cover near the land/ sea boundary spectral indices will have different impact on the identification capacity of the algorithm. Indices like NDWI [16] and MNDWI [17] have been successfully tested for water bodies detection.

For instance, the ROI for the purposes of tests with NDVI-, BNDVI-, GNDVI-based binary classification and waterline production for the Barcelona case study site included sandy beaches interrupted by manmade structures next to the Port of Barcelona. This allowed for testing the processor for different land-sea interfaces.

[15] Ryu, J.-H., Won, J.-S. & Min, K. D. Waterline extraction from Landsat TM data in a tidal flat: A case study in Gomso Bay, Korea. *Remote Sensing of Environment* **83**, 442–456 (2002).

[16] McFEETERS, S. K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing* **17**, 1425–1432 (1996)

[17] Xu, H. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing* **27**, 3025–3033 (2006).

The comparison of waterlines from a scene in Barcelona with slight water disturbances shows stark differences in the performances of the different band ratio-based processes identifying the instantaneous water/land interface, with the with the BNDVI outperforming the GNDVI and NDVI bands in this case (Figure 2.7). The water disturbances appear to increase variation in the red and green pixel values, resulting in the NDVI and GNDVI processors identifying land/sea interface in the sea.

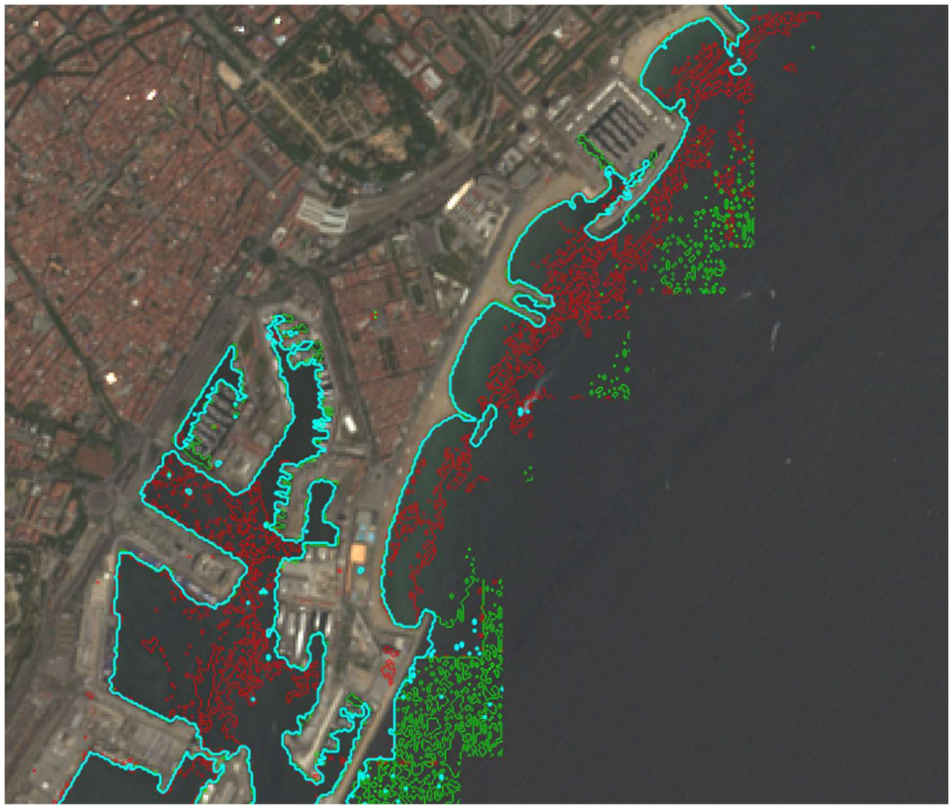


Figure 2.7: Waterline performance differences for different spectral indices: NDVI (red), GNDVI (green), BNDVI (cyan). Barcelona (tile: 31TDF), 02/07/2019.

2.3.4 *Error estimation and uncertainty*

Estimating errors in a statistically robust manner will be constrained by availability bias stemming from the effects of clouds and other weather-related phenomena regarding optical EO products. Waterlines from the waterline processor algorithm feature a certain level of positional uncertainty as well. As with any EO-based product, a number of general remarks can be made regarding these components of the error term.

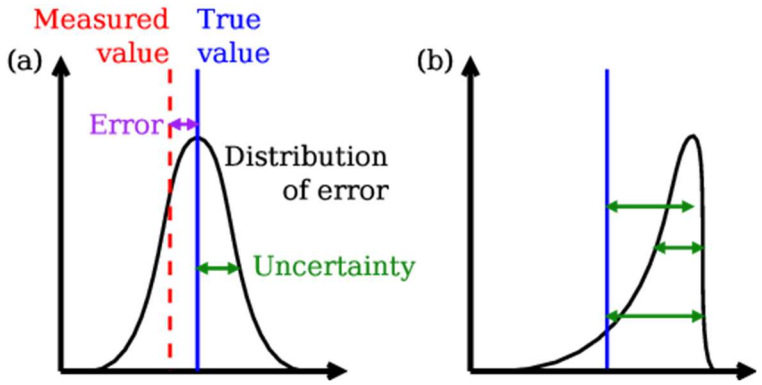


Figure 2.8: Illustration of error and uncertainty. Error in measurement (purple), uncertainty (green), (a) conventional random error, (b) systematic component error, cannot be characterized with a single value (Povey and Grainger, 2015)

As per section 2.3.2, the waterline was defined as:

$$WL^{\mathrm{EO}}(t)=\left\{\left(\vec{X}^{\mathrm{EO}}_{\mathrm{WL,h}},\delta z^{ext}_{WL}\right)(t,s)\middle|s\in\left[s_0(t),s_1(t)\right]\right\}$$

If *t* is the instantaneous time of the snapshot:

- the satellite sensor may take a few seconds to pick-up the datasets that build up an image, e.g. Sentinel-2/MSI has three arrays of sensors that are not in the same focal plane, and look at different areas of the earth, leading to a temporal adjustment from one array to the other in order to deliver a consistent image;
- the timescale of the variability of the waterline is in-between a few seconds to 1 hour, approximately.

Additionally to the length-scale of the observations given by the EO sensor resolutions (10 m for S2/MSI), the length-scales of the phenomenon, which build-up the variability of the waterlines, need to be considerate namely, i. the geomorphological length scales, and ii. the hydrodynamic length scales / or combined hydro-morph length scales which go down to a few decimetres.

Sources of uncertainty can be classified in two distinct categories; radiometric, that is, related to the spectral values obtained by the sensor and subsequent processing, and positional, specifically the accuracy and precision of image geolocation and co-registration, with potential errors that could arise as a result of this. Possible error may be classified in different categories; errors from the instrument, the satellite sensors, errors from product pre-processing like co-registration or geometric and radiometric correction, and errors from the waterline extraction model such as the accuracy of computed band ratios and quality control scores.

The signal to noise ratio (SNR) is a key parameter of satellite sensors, it quantifies how much the signal has been corrupted by noise. It characterizes the actual information content in an image – radiometric resolution. The radiometric resolution of an imaging system describes its ability to discriminate very slight differences in energy.

2.4 Algorithm output

Proxy-based and datum-based shorelines are delivered with their metadata, building on the waterline product which is the output of the described algorithm. The waterlines are delivered in vector format and will be compatible with common GIS software. See more details on it in the TSD (Ref: SO-TR-ARG-003-055-009-TSD, AD-7) in table 5.2.3.

2.4.1 *Product content and organisation*

Waterline products are continuous vector multi-lines linked with their metadata. They are derived using different spectral properties of optical data sets and exploiting data at different resolutions. The product is the instantaneous line at the boundary of the sea and land, providing an instantaneous shoreline proxy. It is computed using a locally adaptive thresholding method on spectral indices created based on the EO products: NDVI, GNDVI, etc. according to the specificities of the region of interest, based on the experience of the operator. The vector multi-line GeoJSON file contains internal quality scoring and will be organized by region in a web-based Data Portal (DAP), along with their metadata for partners to access.

2.4.2 *Quality Control*

There are two approaches to quality control (QC), internal and external.

The internal QC analyses the geometric characteristics of the waterline in segments. For a given segment, it looks at three geometric characteristics in particular;

- the line compactness index (LCI) compares the shape to a cercle, the closer the shape is too a circle the higher is the compactness index.

$$LCI = \frac{4\,\pi\, Area_{convex}}{(Perimeter_{convex})^2}$$

- the line length (LL), measure the geometrical length of a segment. The line length index returns 100 for line longer than 5km and $\frac{lenght}{50}$ for smaller line. Long continuous waterline segments have higher quality flag number and flag very small waterline segments often detected due to breaking waves or landside bodies of water;

- line eccentricity index (LEI) for closed segments. A segment is considered closed if the first coordinate of the geometry is equivalent to the last coordinate, else it is open.

$$LRI = \frac{Length_{shor\ \ axis}}{Lenght_{long\ axis}}$$

LEI shows good ability to isolate the majority of small errors and provide high QC score to short accurate waterline segments. However, this method struggles with complex coastal geometries such as rocky headlands, harbors, groins, and river outlets.

- line rugosity index (LRI) for open segments:

$$LRI = \frac{(X_{max} - X_{min}) + (Y_{max} - Y_{min})}{L}$$

with (Xmas, Ymax) the coordinate of the final point and (Xmin, Ymin) the coordinate of the first point.

L is the total length of the line

QC_intern blend the different indicators scores and ensures that drawbacks from each method are mitigated, this allows for more accurate flagging of both accurate and erroneous waterline segments. The internal QC score is assigned to each segment, 100 being very high confidence that the segment of waterline is the correct line as it appears in the satellite image, and 0 being very low confidence.

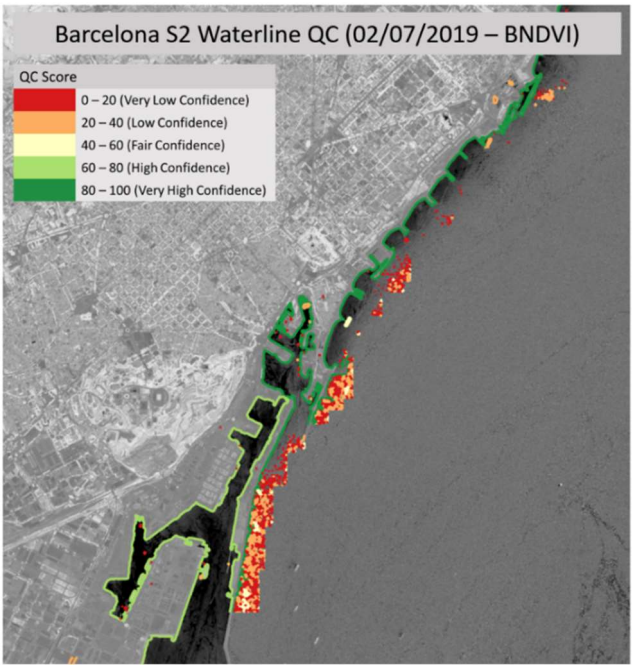


Figure 2.9: Internal quality control, noise correctly identified by lower scores (red and orange).

The external quality control approach looks at waterlines in series and assesses the repeatability of the waterlines. It uses a heatmap approach. The concept applied to waterlines generates a score based on their repeatability, or conversely their randomness. Lines which follow the trend of the other waterlines in series are assigned a high score (maximum=100).

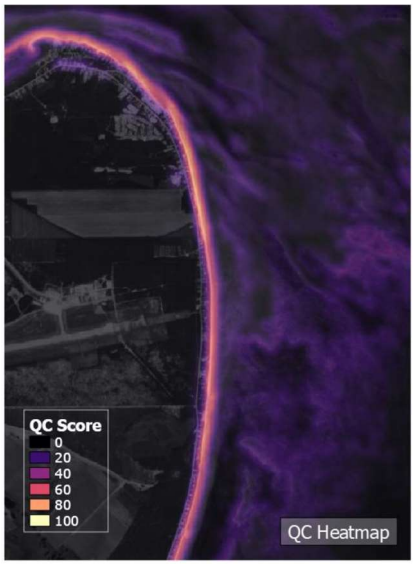


Figure 2.10: External quality control – darker colours representing noise, and the bright line around the coast showing the density of waterlines where we would expect to detect them.

2.5 Algorithm Performance Estimates

2.5.1 *Test specification*

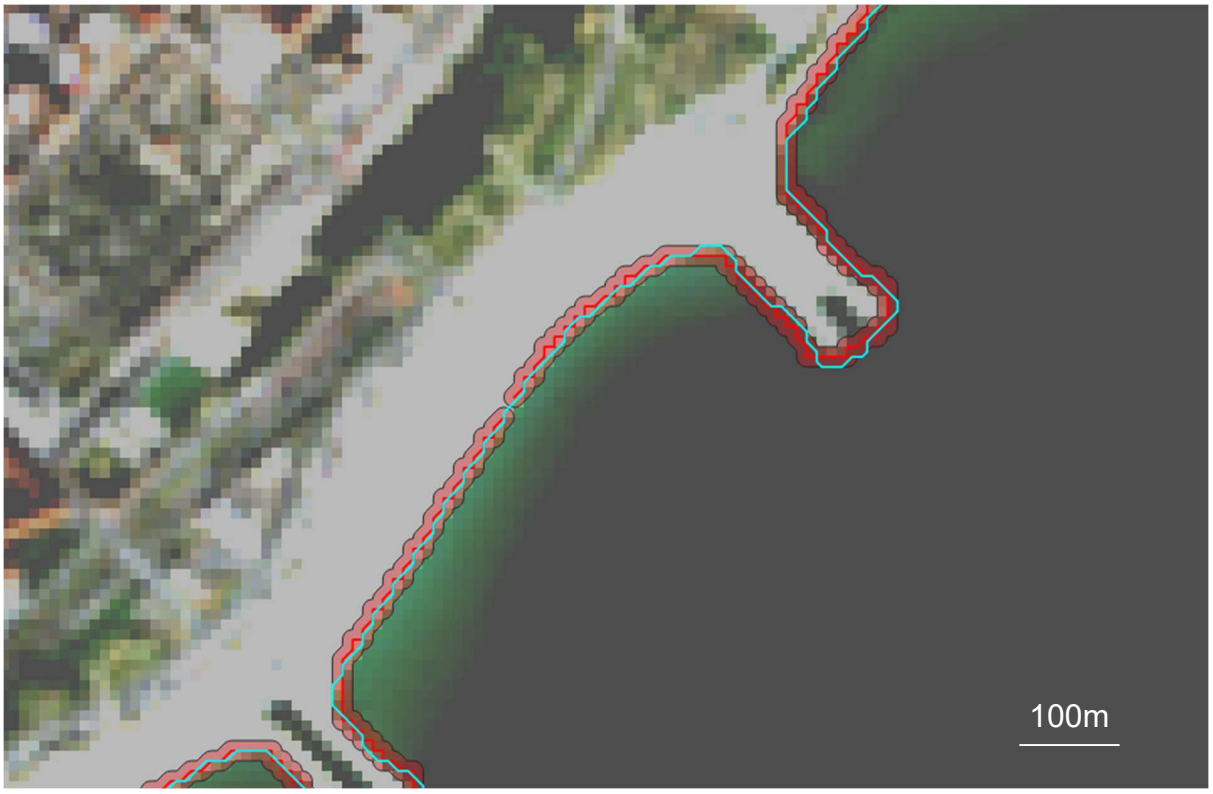
Waterlines created using the current processor (V3 waterlines) were compared to waterlines generated using the previous version from the same images (V2 waterlines) that had been validated and accepted by end users. Comparing the similarity of these waterlines would be a good indication of the accuracy of the V3 waterlines. These tests were conducted internally by producers of the waterline processor.

2.5.2 *As well as a qualitative visual comparison of waterlines, a quantitative approach was taken using our lines comparison algorithm that calculate the percentage of line contained within an increasing buffer around a reference line. This process was repeated, adding 1m to the buffer with each iteration until the total distance of the buffer was 20m. With each iteration, the percentage increase of contained line from the last buffer was also calculated.Test datasets*

33 and 56 V2 and V3 waterline pairs from single Sentinel-2 images were chosen at random for sites Start Bay and Barcelona. The dates ranged from May 2017-October 2019 for Barcelona and May 2017-November 2020 (Start Bay) and were evenly spread out across the years to account for any seasonal variations. These two sites were chosen as between them, they contain a range of waterline features such as smooth beaches, rough coastline and man-made structures (e.g. ports and harbours).

2.5.3 *Validation*

The V3 waterlines of Barcelona and Start Bay were found to be very similar to the V2 waterlines in terms of coverage, even for complex and rugged coastline in Start Bay (Figure 2.11), though white water decreased the similarity.



(a)

(b)

Figure 2.11: V3 waterline (cyan) against 10m buffer around V2 waterline (red) for a section of (a) Barcelona and (b) rugged coastline in Start Bay.

There is significant dissimilarity of the waterlines in coverage of complex areas such as ports, however visual inspection indicated that this is due to an increase in accuracy in new waterline in such areas (Figure 2.12).

Figure 2.12: V3 waterline (cyan) and V2 waterline (red) for a port in Barcelona

The outputted waterline using the V3 processor has also been found to be far neater than the output of the previous version. A BNDVI-based V3 waterline saw vast improvements compared to the V2 processor waterline for the same tile on the same day that was largely affected by artefacts in the image (Figure 2.13).

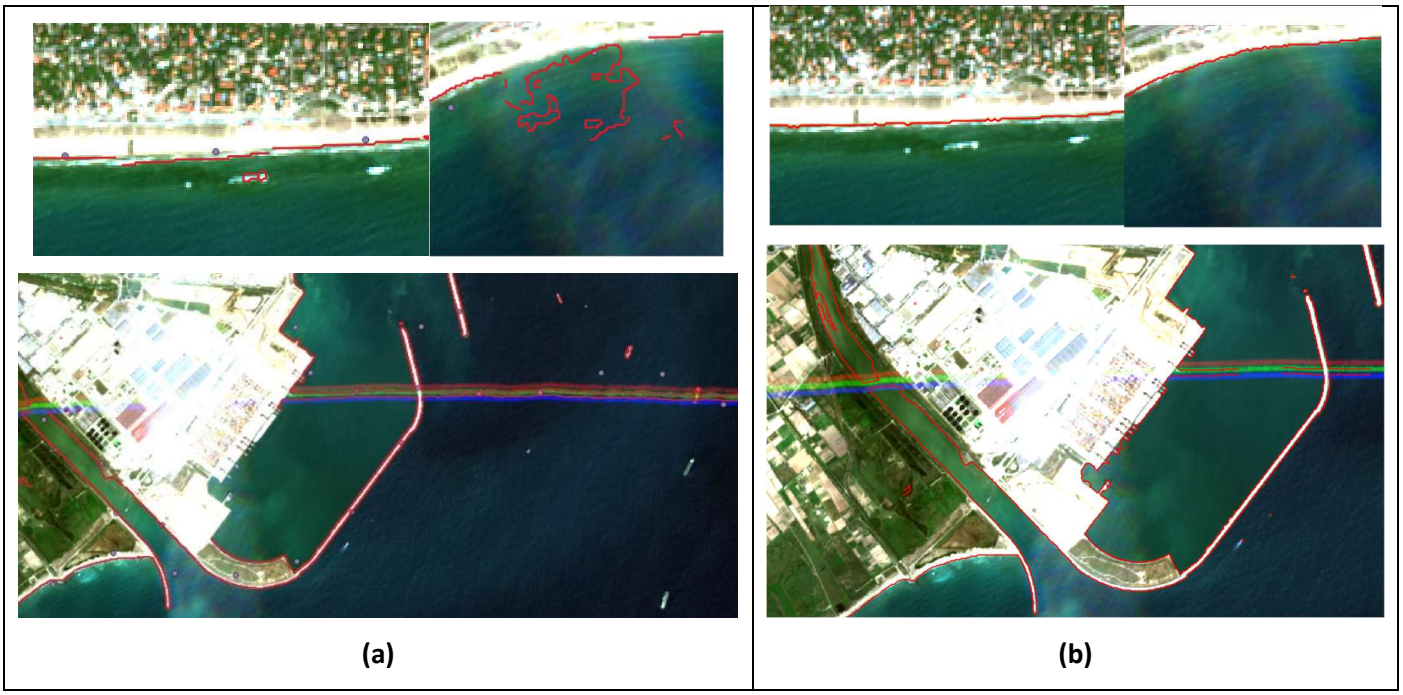


Figure 2.13: Clouds, white water and passing airplane causing discontinuities and false edges in BNDVI-based (a) V2 waterline processor compared to (b) BNDVI-based V3 waterline processor, Barcelona, 21/08/2019 S2 tile T31TDF

3 Conclusion

As stated, the objective of the Coastal Erosion from Space project and the corresponding CCN is to retrieve Coastal State Indicators (CSI) describing the dynamics and evolution of coastal systems. After a year and a first version of the waterline processor, the end-user community validated the V2 waterline and highlighted some possible improvement.

This new version of the waterline ATBD describe the improvement of the waterline processor allowing us to obtain more continuous lines without the pixel square shape using the marching square vectorisation. The validation section demonstrate how those updated lines behave in various coastal environment and their similarity, in term of spatial accuracy, with V2 waterlines.

Waterlines are being produced for the past 25-year for all the study sites, for all available EO observations allowing to monitor local erosion and the occurrence of significant events, contributing to our understanding of coastal processes and to a quantitively sound assessment of erosion and accretion processes in the short and long term. Partners will access the waterlines as GeoJSON files together with their respective metadata files.

3.1 Assessment of limitations

The limitations identified to date stem from the method used: edge detection. After a binary classification of the image to either land or water, the algorithm creates a line at the boundary of the two. However, this line is not always continuous, and the boundary classification is not always correct. Therefore, the resulting discontinuities and false edges pose problems for being able to define the final waterline product in certain cases. This is particularly the case in complex estuarine environments (sediment load); however, issues are present at less complex coastal areas with very transparent waters as well (nearshore bathymetry). White water (WW) is an additional concern, especially in VHR imagery. The spectral response from crashing waves is dramatically higher than calm water and as such the processor identifies it as land and defines a threshold seaward of the waves. The increased albedo and reduced absorption results in significantly higher NIR values, which in turn produce values in the band ratios that are somewhere in between common peaks for land and sea.

The issues of the jigsaw effect and discontinuity of the previous version of the processor have largely been overcome with version 3, though a new issue arises of long waterline segments being in one part very accurate and the other very noisy. The new waterline processor also reduces the noise generated from crashing waves near the waterline, however, white water has still been found to reduce accuracy.

3.2 Mitigation

Most limitations identified from phase 1 have been mitigated using:

- Pre-processing testing for the optimal band ratio – this can improve the overall quality of the waterline, mitigating against false edges and noise. It depends upon the physical characteristics of the co-registered image, e.g. suspended sediment levels, gradient of slope, and white water cusps.
- Inclusion of a rough threshold value – this assists mitigation of false edges and fragmented delineation. Where spectrally similar signals produce two troughs on the histogram of values, the rough threshold aids the selection of the 'correct' value, improving continuity in the waterline delineation.
- Internal and external quality control assigns a confidence level to the data and therefore allows data cleaning, mitigating against the use of erroneous data.
- Optimal kernel size of 75 x 75 pixels for sentinel-2 processing aids reduction of false edges from white water

4 References

Alesheikh, A. A., Ghorbanali, A. & Nouri, N. Coastline change detection using remote sensing. Int. J. Environ. Sci. Technol. 4, 61–66 (2007).

Dolgopolova, E. N. & Isupova, M. V. Water and Sediment Dynamics at Saint Lawrence River Mouth, ISSN 0097-8078, Water Resources, 2011, Vol. 38, No. 4, pp. 453–469. © Pleiades Publishing, Ltd., 2011.

Frazier, P. S., Page, K. J. Water body detection and delineation with Landsat TM data. Photogrammetric Engineering & Remote Sensing vol. 66 1461–1467 (2000).

Keuchel, J., Naumann, S., Heiler, M. & Siegmund, A. Automatic land cover analysis for Tenerife by supervised classification using remotely sensed data. Remote Sensing of Environment 86, 530–541 (2003).

Kumer Ghosh, M. et al. Monitoring the coastline change of Hatiya Island in Bangladesh using remote sensing techniques, ISPRS Journal of Photogrammetry and Remote Sensing, 137–144 (2015).

Lillesand, T., Kiefer, R. W. & Chipman, J. *Remote Sensing and Image Interpretation*. (John Wiley & Sons, 2004).

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

Maglione, P. et al. Coastline extraction using high resolution WorldView-2 satellite imagery, European Journal of Remote Sensing, 47:1, 685-699 (2014).

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

Mohammady, M., Moradi, H. R., Zeinivand, H. & Temme, A. J. A. M. A comparison of supervised, unsupervised and synthetic land use classification methods in the north of Iran. Int. J. Environ. Sci. Technol. 12, 1515–1526 (2015).

Muttitanon, W. & author, N. K. T. C. Land use/land cover changes in the coastal zone of Ban Don Bay, Thailand using Landsat 5 TM data. *International Journal of Remote Sensing* **26**, 2311–2323 (2005).

Ryu, J.-H., Won, J.-S. & Min, K. D. Waterline extraction from Landsat TM data in a tidal flat: A case study in Gomso Bay, Korea. *Remote Sensing of Environment* **83**, 442–456 (2002).

Stafford, D. B. Air Photo Survey of Coastal Erosion. Photogrammetric Engineering & Remote Sensing 11 (1971).

Warmerdam, F. & Rouault, E. a. o. gdal_contour. GDAL documentation, [Online] Available at: https://gdal.org/programs/gdal_contour.html (2022) [Accessed 21 February 2022].

Xu, H. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing* **27**, 3025–3033 (2006).

End of Document