Coastal Erosion from Space

Coastal Erosion time series in Happisburgh (GB)

Coastal Erosion events:
in Nerja (Spain) in Kilkee (Ireland) in St Laurent mouth (Canada)

Algorithm Theoretical Baseline Document

Ref: SO-TR-ARG-003-055-009-ATBD-SF
Date: 30/10/2019

Customer: ESA

Contract Ref.: 4000126603/19/I-LG

This page is intentionally left blank

Version history

Version	Date	Modification
Version 1	09/12/19	Review
Version 2	20/01/2021	Final review
Verification by	François-Regis Martin-Lauzer	
Authorisation	Craig Jacobs	

Acronyms

CESBIO : Centre d'études spatiales de la BIOsphère (Fr)

CNES : Centre national d'études spatiales (Fr)

DSI: Datum-based shoreline indicators

DTM: Digital elevation model

EO: Earth observation

GRD: Ground range detected

HR: High resolution

LiDAR: Light detection and ranging

LL: Littoral line

LULC: Land use land cover

NDSI: Normalized snow index

NDVI: Normalized difference vegetation index

NDWI: Normalized difference water index

NIR: Near infra-red

PSI: Proxy based shoreline indicators

RF: Random forest

SAR: Synthetic-aperture radar

SfL: Seafront Line

SLC: Single look complex

SVM: Support vector machine

SWI: Superfine water index

TIR: Thermal infra-red

TOA: Top of atmosphere

URD: User requirement document

VHR: Very high resolution

VNIR: visible and near infra-red

Applicable and reference documents

Id	Description	Reference
AD-1	Requirement Baseline Document	SO-RP-ARG-003-055-006-RBD_v1.0_20190916
AD-2	User Requirement Document	TR_CR_19_055
AD-3	Geolocation ATBD	SO-TR-ARG-003-055-009-ATBD-GL

Content

List of Figures

1 Overview and Background Information

1.1 Product requirement

According to partners and end-users' requirements listed in the URD regarding shore management, the seafront processor aims to compute indicators linked with coastal environments. Coastal indicators, due to their ability to determine changes to the extent of the littoral zone are key for our understanding of coastal erosion.

Figure 1.1: Example of a range of visibly discernible shoreline indicator features and their spatial relashionship, Duranbah Beach, New South Wales, Australia, © Boak & Turner, 2005

Figure 1.2: Littoral line (S), seafront (E), dry sand/wet sand boundary (K) and debris line (G) from a 3d representation of Duranbah Beach, New South Wales, Australia

1.1.1 Information content quality and value

The landward boundary of the littoral zone (Littoral Limit) is characterized by a vegetation line, rocky escarpment or a top cliff in semi-natural environments and a sea wall or other anthropogenic structures in urban areas. Numerous scientific articles deal with the coastal zone, its studies, its management. However, in those studies, no consensus emerges regarding to the definition of the limit of the shoreline or the backshore areas. The littoral limit is the line at the boundary between what is commonly called beach and the rest of the land. On-site, this delimitation is easily recognized, either by the topography (beach elevation profile is linear, no breaks or steps in the evolution of the elevation) or by the type of soil (without any breaks a change of land cover, from sand to rocks, or from gravels to vegetation). For convenience, we accept temporary build-up or small built-up areas as part of the beach (summer straw huts, beach restaurant). In some cases, an intermediate area can be observed in case of mudflat, saltmarsh, dunes, etc. In this case, the littoral limit is the boundary between the backshore and the intermediate area, and the delimitation between the intermediate area and the beach is the seafront line.

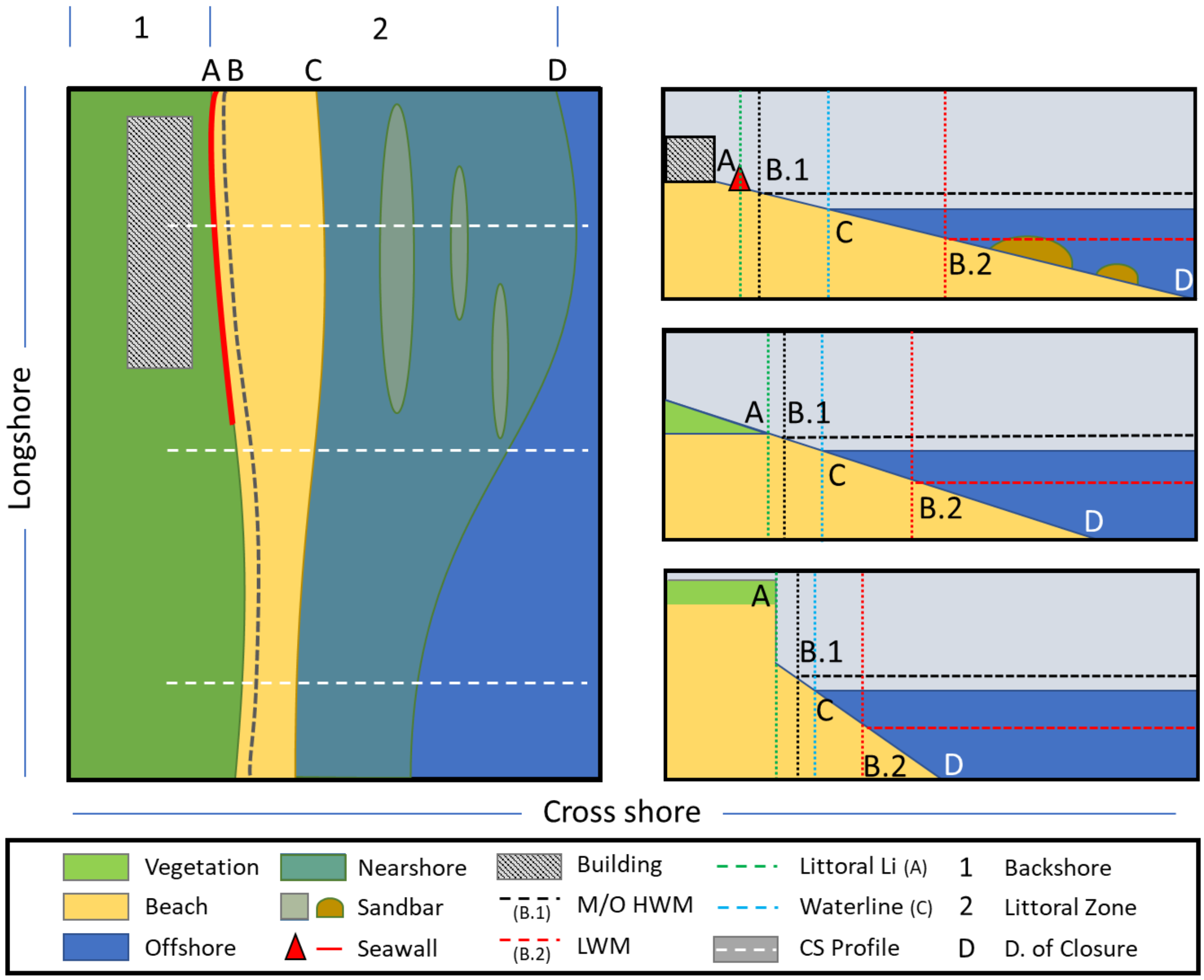


Figure 1.3: Shore and coastal features diagram

To monitor and study shore evolution, erosion rate and sediment transportation, we look at the changes in the Littoral line (A in figure 1.4) position.

For the production/extraction of the Littoral and Seafront lines, an intermediate product is produced. LL and SfL are derived/deduced from a Land Use Land Cover (LULC) map. Production of an annual LULC map will give us the different type of soil cover of our ROI, i.e. the coast area, thanks to land cover information, we define what type of land cover belongs to the backshore and which one to the littoral zone. Land cover identification depend on the satellite image resolution and photo interpretation.

Image resolution, very high resolution (VHR) imagery around 2 or 5 m of spatial resolution and high resolution (HR) imagery around 10 or 30 m are useful in different contexts of coastal erosion. Erosion regimes are not equal along the Earth coastlines, not even along a 500 m beach. Some areas undergo changes in the order of several tens of meters each month while for others, changes are centimetric. To be able to monitor those changes with the most accurately possible, different resolutions will be used.

1.1.2 Product order and delivery services

1D littoral lines and 2D LULC map products, will be produced annually even maybe seasonally from HR data, for all sites for the past 25 years.

Output format will be a vector ESRI shapefile format for the Littoral and Seafront lines, and a raster GEOTIFF format for the LULC maps. Both being compatible with GIS software as ArcGIS and QGIS.

Products will be available on a geoportal with their quality control information and delivery will be available by an ftp transfer or by using the web service developed for the project.

1.2 Feasibility review

Classification maps are widely used in land use studies, they allow one to represent at a given moment, soil conditions and their evolution. Thanks to the number of EO satellites (revisit time) and the variety in their specification (spatial resolution) there are a number of different land cover map available.

1.2.1 Satellite sensors and mission

Please refer to the pre-processing ATDB (ref: SO-TR-ARG-003-055-009-ATBD-GL)

1.2.2 Existing EO Products

Please refer to the pre-processing ATDB (ref: SO-TR-ARG-003-055-009-ATBD-GL)

1.2.3 Models specifications

Image classification methods can be divided into unsupervised and supervised classification methods. Supervised classification, , use prior knowledge which improves the classification accuracy[1]. In previous

---

[1] 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).

studies, supervised methods were evaluated, different measurements for performance estimation were tested to compare different classification algorithms. Complex measures based on the combination of other measures are difficult to interpret, thus, chosen simple measures will allow a straightforward interpretation[2]. Supervised classification will be used for our process to ensure a greater accuracy, however there is different algorithm for supervised classification with their own specificities and weakness.

Comparative analysis of different classification algorithms was conducted by comparing their robustness to parameter changes or their accuracy. Classification accuracy depends on various parameters, number of sub-classes and other parameters inherent to the algorithm[3] such as decision rules.

In addition to the classifier type, variety and parametrisation, image resolution also impacts the classification result. From VHR to low-resolution image the size of the smallest distinguishable unit increases. The influence of image resolution on the classification relies on spatial inter-class variability and intra-class variability. Thus, we may have an optimum resolution by land cover above and below which classification accuracies will decline[4]. The optimum resolution for classification strongly depends on the landscape. The more fragmented and mixed the landscape, the finer the resolution should be chosen[5].

Two classification maps, based on the same classes and classes' definition, may vary although the objects have not changed, because the objects' inherent or apparent optical (VNIR-TIR) properties have changed through time (seasonal change, agricultural cycles). Multi-temporal classification has proved effective for land cover type discrimination[6]. However, the use of multi-temporal data does not always ensure an increase in accuracy. The realization of a temporal classification map requires a real knowledge of land cover. Features gathered in the same class need to evolve in a similar way. Low intra-class variability is the key for accurate classification, i.e. good training data set. Building such a training set requires time and knowledge, thus a balance needs to

[2] Amancio, D. R. *et al.* A Systematic Comparison of Supervised Classifiers. *PLOS ONE* **9**, e94137 (2014).

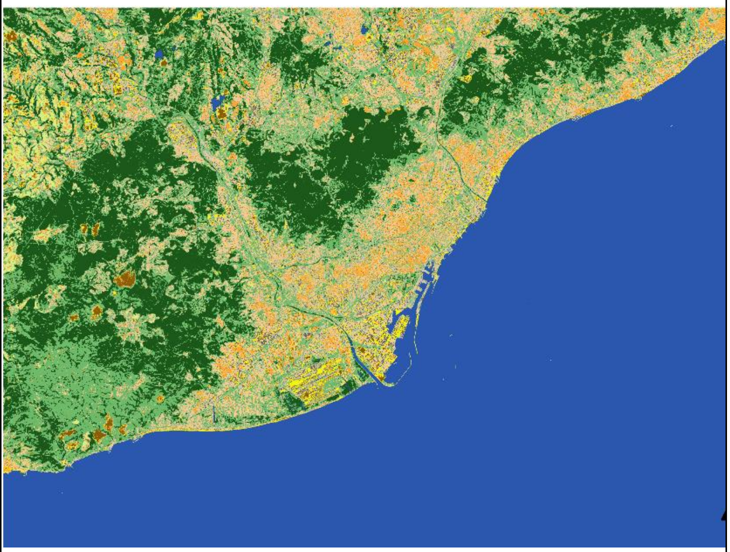
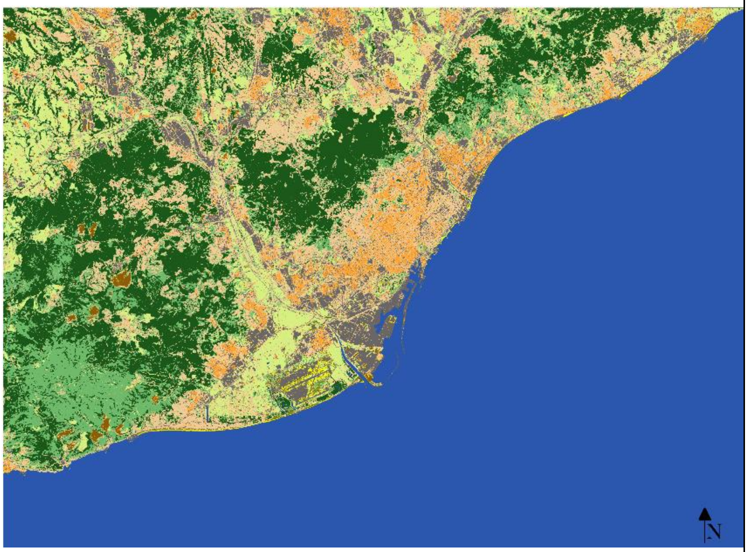
[3] Keuchel, J., Naumann, S., Heiler, M. & Siegmund, A. Automatic land cover analysis for Tenerife by supervised classification using remotely sensed data. *Remote Sensing of Environment* **86**, 530–541 (2003).

[4] Townshend, J. & Justice, C. Information extraction from remotely sensed data. Remote Sensing (2007)

[5] D. Chen, D. A. Stow & P. Gong. Examining the effect of spatial resolution and texture window size on classification accuracy: an urban environment case. *International Journal of Remote Sensing* **25**, 2177–2192 (2004)

[6] Langley, S. K., Cheshire, H. M. & Humes, K. S. A comparison of single date and multitemporal satellite image classifications in a semi-arid grassland. *Journal of Arid Environments* **49**, 401–411 (2001).

be found between available resources and required accuracy. Moreover, temporal changes are strongly correlated with microclimate, thereby new data sets may be needed for the different study areas.

Random forest from a single image	**Random forest from 12 images**
	
Overall accuracy: 0.663, KAPPA coefficient: 0.626	Overall accuracy: 0.926, KAPPA coefficient: 0.917
 Confidence 	 Confidence 

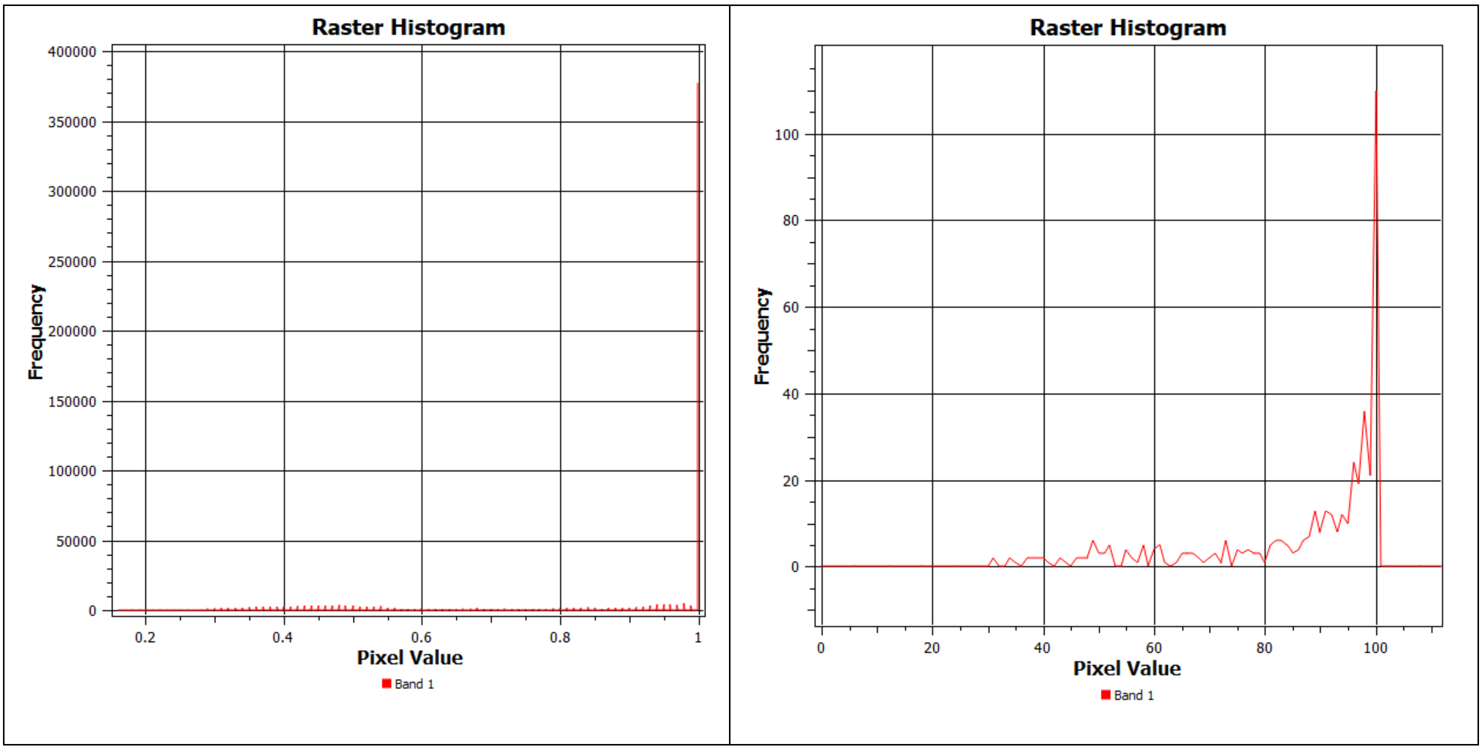


Figure 1.4: Comparison of a RF classification map with 1 single image and with a time series of 12 images

Lines are extracted from a LULC map computed from a supervised classification. Iota² (Infrastructure pour occupation des sols par Traitement Automatique Incorporant les Orfeo Toolbox Applications) is a satellite imagery processing chain built to generate land use and land cover maps over entire countries. Iota² can manage a large volume of data over large areas: it is notably used yearly to generate the OSO (Occupation des Sols) map in France since 2009. One of the main draws of Iota² processing chain is that, unlike most semi-automatic classification chains that use a single image to produce the classification, iota² is able to process a time series of images, independent of cloud cover (provided a cloud mask is included in the entry data). Classes chosen by operators are characterized by iota by the succession of spectral information. Training and classification are not done with one image but with a time series, consequently the classes are defined by the variation of the reflectance of the land cover from one image to another. Consequently, capturing agricultural cycle and seasonal changes, however punctual event like fire and flooding will appear as an anomaly in the classification result.

1.2.4 Auxiliary data

LL and SfL will be deduced from land occupation; thus, an accurate classification map is needed to extract the position of those lines. If unsupervised classification methods don't require prior knowledge, supervised classification ones need a training data set as input which contains labels for some sample locations. To improve the learning phase, scientists use DTM or/and texture information.

Validation data will be needed to check the accuracy of our detected line. Those data may come from an eye comparison between our line and VHR satellite data, or from LiDAR surveys.

1.2.5 Currently known issues

Known issues are:

- Spatial shift between images
- Radiometric and geometric distortion
- Features identification due to similar spectral properties
- Additional noise in VHR data due to more detailed scene
- Erosion rate is a volume indicator, how we access to the elevation information

First, images from different sensors or taken at different dates may not be spatially aligned. Scene illumination and viewing geometry may vary between two acquisitions.

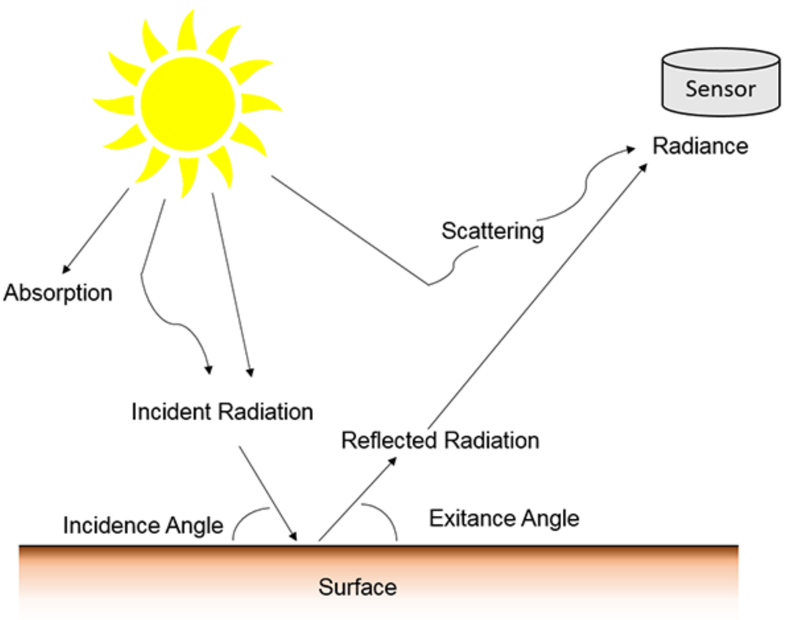


Figure 1.5: Illustration of the various process impacting scene illumination (Geospatial Science at Humboldt State)

To detect the littoral line, we need to identify limits between beach areas and several different types of land cover. In Barcelona land cover adjoining the beach may be built up areas, vegetation, cliff (see figure 1.5).

From previous studies conducted by ARGANS, separation between beaches and built up areas, specifically urban areas may be fuzzy and not detected with high precision due to their spectral similarities.

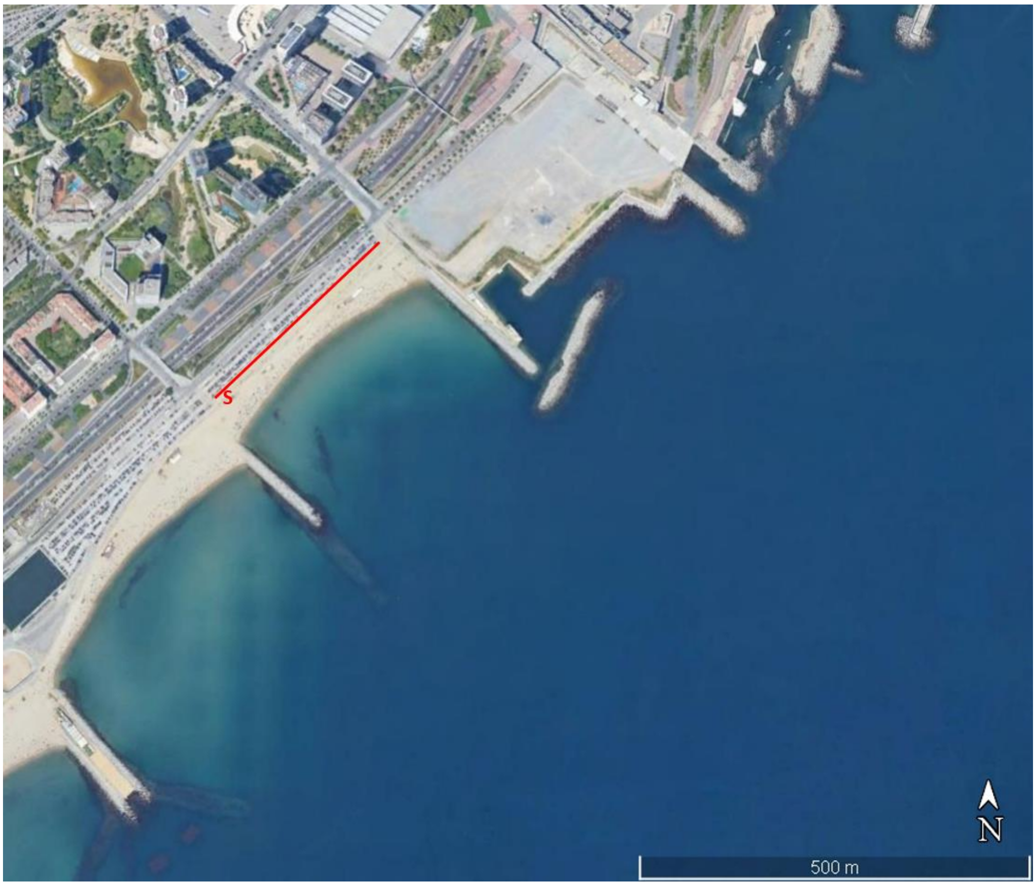


Figure 1.6 :Zoom over a littoral line (s) in Barcelona

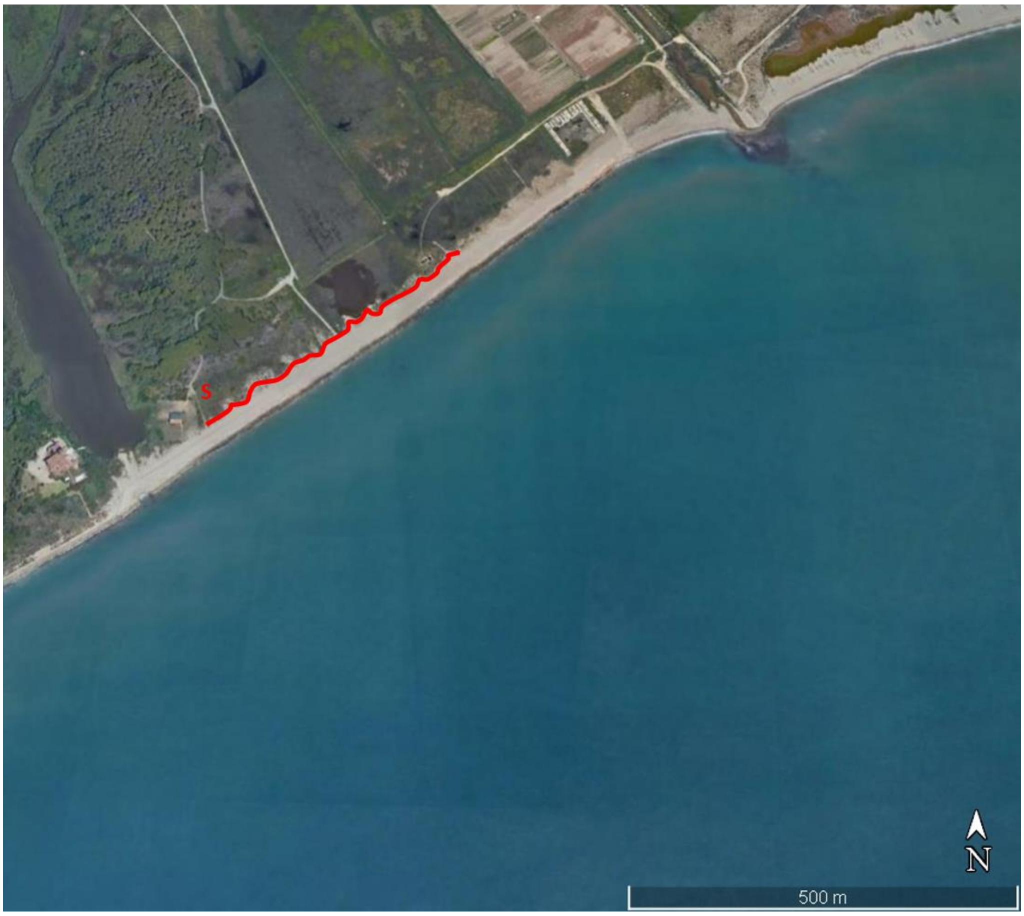


Figure 1.7: Zoom over a littoral line (s) in Barcelona

Classification of VHR data with iota² will raise some issues. Due to iota² time series processing some small un-periodic event usually not visible on HR data may confuse the classification. On the beaches of Barcelona, every summer, vacationers install their towels and umbrella. On a HR image we cannot visibly identify them, information is averaged according to the pixel scale. However, on a VHR image those features are distinguishable (see figure 1.9).

The delimitation of the seafront may be identified thanks to photointerpretation. Operators can identify some features like top of a dunes or the end of it (see figure 1.3) even if there is no real visual differences. However, finding this boundary automatically may be difficult as the seafront is mainly defined according to a height difference.

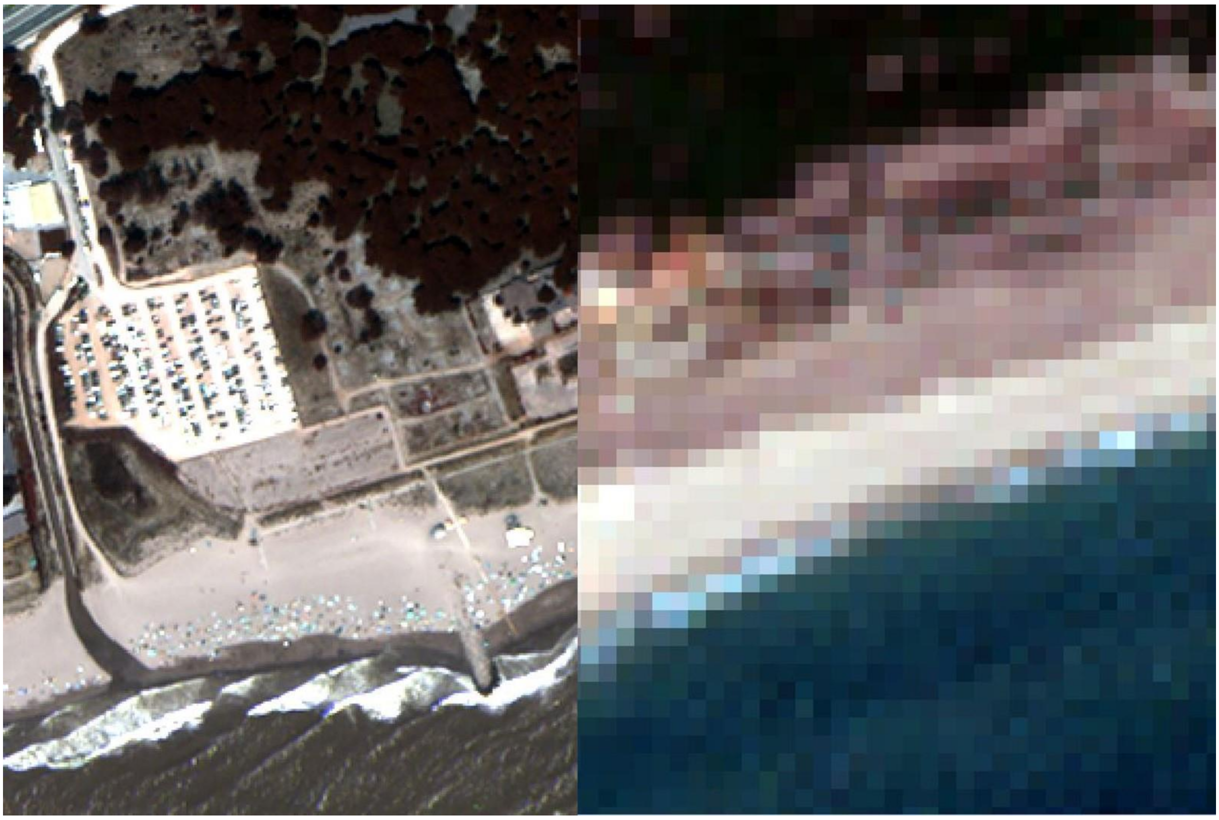


Figure 1.8: Comparison of an WorldView 2 image (left) and a Sentinel-2 image (right) over a beach in Barcelona in July 2018

1.3 Potential Solutions

Radiometric and geometric variations are well known issues, most are corrected during imagery pre-processing –Sentinel-2 level 1B product is radiometrically corrected and physical geometry is refined. Most of pre-processing also provide orthorectified, TOA reflectance corrected products.

All satellite data used in input of the classification processing chain have been co-registered on a reference product with a high spatial accuracy. All data are thus perfectly overlapping.

Our main focus with the littoral line is to detect and monitor their evolution in response to beach sediment volume changes. When a beach is joining a built-up area, a wall is most of the time build to stop beach progression inland. Littoral lines, in this blocked environment, are relatively stable in opposition to littoral line between beach and vegetation. In that case, littoral lines recede or advance according to the beach sediment movement. Most modifications will be thereby visible between beach and vegetation. Thus, we may focus and improving the processor for those areas.

1.4 Product Specifications

The process will provide littoral and seafront lines product. During the process an intermediate 2D product will be computed: the land use land cover map. Those products will be computed annually and seasonally (if EO data are available) from a time series analysis. Products in vector or raster format will be available on a geoportal and delivered by ftp transfer, they will be compatible with GIS software such as ArcGIS or QGIS.

Sentinel-2 and Landsat data will be used. VHR data may be used following the erosion regime, for changes less than 10m HR data may not be sufficient. However, the use of VHR data raises some difficulties as more details are visible and will interfere with the classification by adding noise.

Temporal classification will provide the annual LULC map, using a random forest algorithm within the iota² classification processing chain. A specific training data set will be required as temporal changes are strongly linked with the micro-climate of our study areas.

2 Algorithm Description

2.1 Data Processing outline

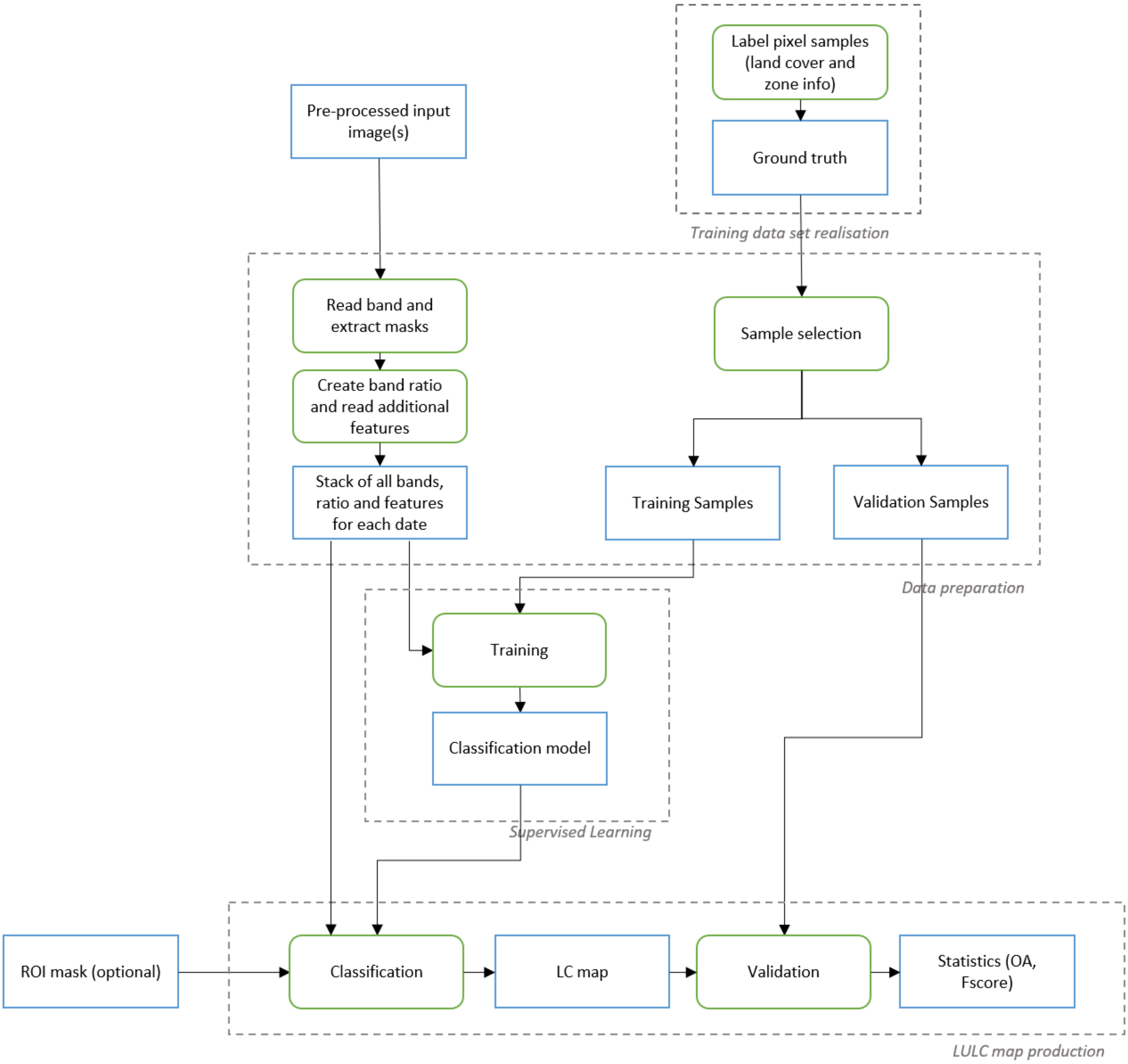


Figure 2.1: Classification processing map

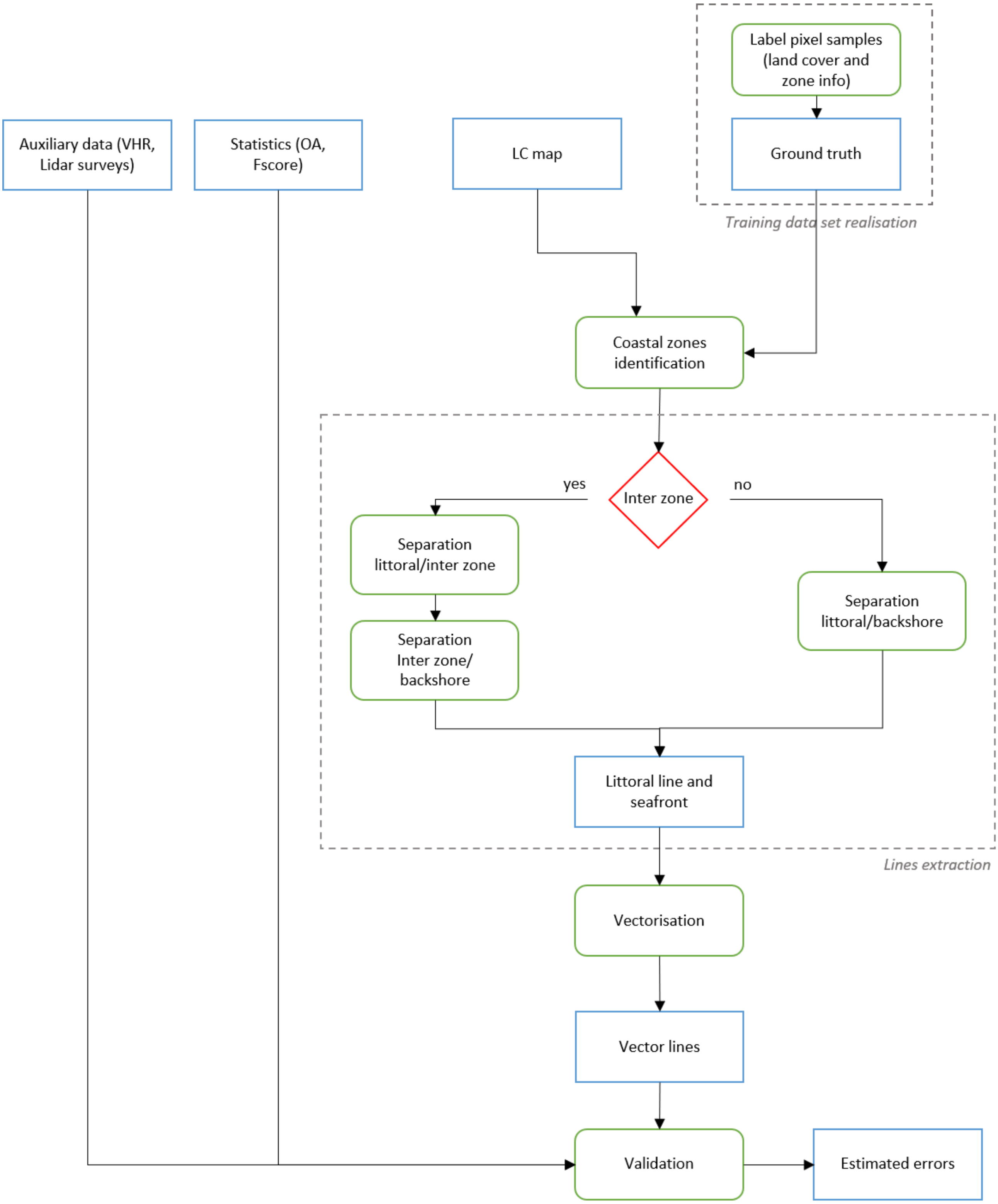


Figure 2.2: Lines extraction processing map

2.1.1 Sketch of the computer program

Land cover identification will be first realised by eye to build a training set to adequately describe the land cover.

First step of the process will be the computation of the validity masks from cloud cover map and land using quality band and mask from the S2L2 classification band and Landsat quality masks. Then samples are extracted and split in two parts for training and validation. Training sets with reference image and features information are used to build the classification model and the validation set is used to get the validation statics from the classification obtained.

Lines will be extracted from this classification map according to the label indicated in the training set.

2.1.2 Pre-requisite

Reference optical images need to be radiometrically corrected. As we work with time series and will realise a comparative study, all images need to be co-registered from a master image.

The training data set needs to be specific to the study area, a small geographical displacement may cause a great inaccuracy due to the specificities of the evolution of the vegetation corresponding the micro-climate.

While building the training data set, required information are a CODE attribute linked with the land cover label and ROI regional attribute linked with the type of regional ecosystem if we use the multi-regional tool. We distinguish 3 different types of zones: the littoral zone (a), the intermediate zone (b) – between the littoral line and the seafront, and the backshore zone (c). The LL is the boundary between a and c, and (a) and (b) when (b) exist. When we have the seafront line, between (b) and (c), when (b) exist.

2.2 Algorithm Input

The process needs different inputs, linked with section 1.2.4, the only auxiliary data we are using is the training data set.

Other inputs will be optical reference images pre-processed according to section 2.1.2, and extra features like band ratio according to the type of land cover.

2.3 Theoretical Description of the models in background of the procedure

2.3.1 *Physical Description*

Most of land covers are not consistent through time. For example, seasons may completely change a forest physicals and visuals characteristics. Human activities can also lead to temporal changes, even if crops are under seasonal cycles, their nature can be artificially modified by changing the type of seedling. These physical changes cause alteration to the reflectance signature of the surface, allowing detection through multi-spectral analysis. These analyses include single-band reflectance alteration and intra-band difference indices. Moreover, some spectral indices may be used to improve the recognition of land cover.

By improving land detection, the accuracy in the localization of the boundary between the different land cover is improved.

2.3.2 *Mathematical Description and calculation procedures*

From a given time series $I = \{ i_1, i_2, \dots, i_n\}$ of *n* images, *n* ∈ *ℕ,* with m features each and a sample training set $Y = \{y_1, y_2, \dots, y_d\}, d \in \mathbb{N}$, for each sample $y_k,\ k \in [1, d]$ a vector of observation $Z = \{z_1, z_2, \dots, z_j\},\ j = n \times m$ is built from all features from all images. An observation z_i is a spectral information from a band or a band ratio at one date, i.e. one image.

Those $Z_d, d \in \mathbb{N}$ vector, one for each training sample, are labelled thanks to the training set in inputs. With a machine learning process on those *d* samples will allow us to classify the tile.

2.3.3 *Acceptance of the Models*

Spectral indices are widely used to identify land cover type. NDVI is mainly used for vegetation monitoring, from biomass estimation[7] to crop assessment[8] but also for non-vegetation features extraction (water bodies,

[7] González-Alonso, F., Merino-De-Miguel, S., Roldán-Zamarrón, A., García-Gigorro, S. & Cuevas, J. M. Forest biomass estimation through NDVI composites. The role of remotely sensed data to assess Spanish forests as carbon sinks. *International Journal of Remote Sensing* **27**, 5409–5415 (2006)

[8] Wardlow, B. D. & Egbert, S. L. Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the U.S. Central Great Plains. *Remote Sensing of Environment* **112**, 1096–1116 (2008).

rocky areas)[9]. NDWI on the other hand, is sensitive to change in liquid water content of vegetation canopy[10]. NDWI is complementary to NDVI, not a substitute. NDWI may also be used to delineate water bodies by enhanced water's spectral characteristics[11]. Other indices may be used according to the different land cover. Supervised classification is realized using thresholds on spectral indices NDVI[12], NDWI[13], NDSI[14], SWI[15] or on band ratios, and even textures.

The random forest algorithm, through several studies, performed some accurate LULC classifications from hyperspectral data[16] and from multisource data[17]. If the random forest classifier can reach a comparable accuracy as SVM, RF algorithm parametrization is easier and quicker than the ones required for SVM. Furthermore, RF classifier provides the relatives importance of different features during the classification process[18].

2.3.4 Error estimation and uncertainty

The accuracy of localization of each pixel in an image depends on the accuracy of satellites' measurements: position (orbit determination), attitude, sensors' position and orientation within the frame of the satellite, and also on the quality of the processes for all pixels to be in an accurate (*x*,*y*) position on the ground, i.e.

---

[9] Bhandari, A. K., Kumar, A. & Singh, G. K. Feature Extraction using Normalized Difference Vegetation Index (NDVI): A Case Study of Jabalpur City. Procedia Technology 6, 612–621 (2012).

[10] Bo-cai Gao. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment* **58**, 257–266 (1996).

[11] Qiao, C. *et al.* An Adaptive Water Extraction Method from Remote Sensing Image Based on NDWI. *J Indian Soc Remote Sens* **40**, 421–433 (2012).

[12] Normalized Difference Vegetation Index

[13] Normalized Difference Water Index

[14] Normalized Difference Snow Index

[15] Superfine Water Index

[16] Ham, J., Yangchi Chen, Crawford, M. M. & Ghosh, J. Investigation of the random forest framework for classification of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing* **43**, 492–501 (2005)

[17] Gislason, P. O., Benediktsson, J. A. & Sveinsson, J. R. Random Forests for land cover classification. *Pattern Recognition Letters* **27**, 294–300 (2006)

[18] Pal, M. Random forest classifier for remote sensing classification. *International Journal of Remote Sensing* **26**, 217–222 (2005)

corrections for optical distortions from the sensor system, and apparent changes in the position of ground objects caused by the perspective of the sensor view angle and ground terrain (orthorectification). Accuracy of intra-satellite measurements are delivered in level 1C products as optical distortion corrections. Orthorectification is processed for level 2 products.

Uncertainties linked with the classification process are calculated and summarized in the confusion matrix and the validity maps available at the end of the process.

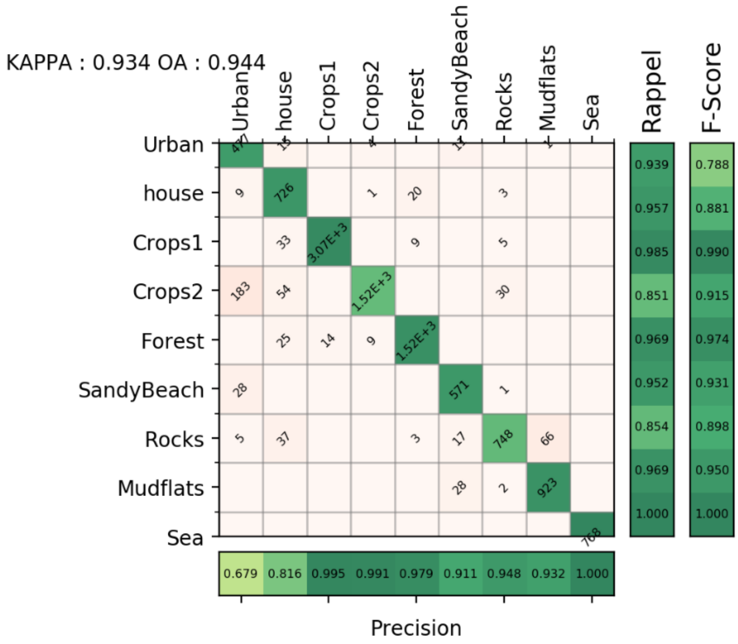


Figure 2.3: Example of a confusion matrix with the recall, precision and F-score metrics

2.4 Algorithm output

Proxy based littoral and seafront lines will be linked with their LULC map and other metadata. Lines will be delivered in vector format and LULC map in a raster format. All will be compatible with common GIS software.

Seafront and littoral lines products will contain continuous vector lines. LULC map from which lines are extracted is a raster type product, linked with metadata such as accuracy and quality values.

All products will be compatible with common GIS software.

Littoral lines products will be organized in different folder according to the study area. For each area partners will access the littoral limit with all metadata, refer to section 2.4.1.

2.5 Algorithm Performance Estimates

2.5.1 Test specification

Ability of the processor to extract lines strongly depends on the accuracy of the classification process and if the process is able to classify the different littoral and backshore area, i.e. littoral zone, intermediate zone and backshore zone.

Several tests need to be conducted using different amounts of images, for different areas to test our process in different environment.

Precision of the classification process is evaluated through a confusion matrix which gives the overall accuracy and the KAPPA coefficient.

2.5.2 Test Datasets

Process will be tested for all study areas. Indeed, each study location is a unique environment with specific characteristic and changes. Process should be tested for cliffs, hilly areas as well as for large sandy beaches.

2.5.3 Verification

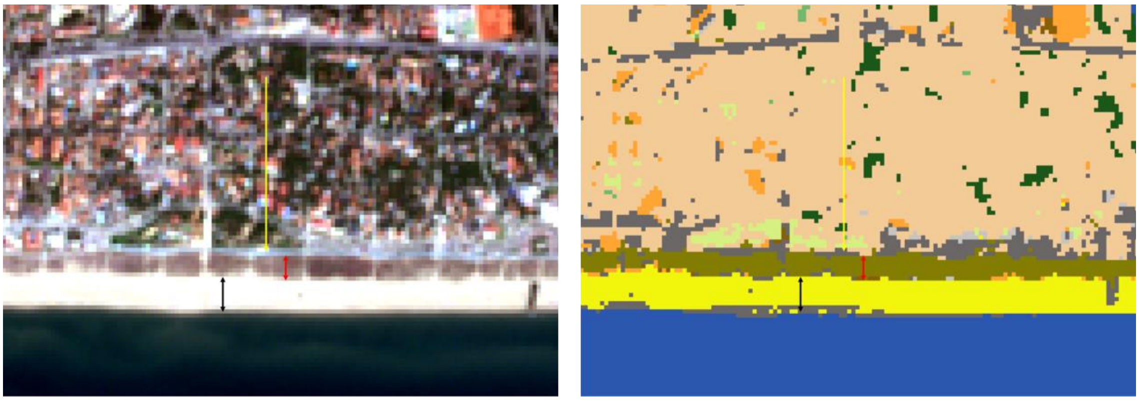


Figure 2.4: Identification of the 3 areas, littoral zone (black), intermediate zone (red) and backshore area (yellow) on both optical and LULC data

On figure 2.4, we can see that the different coastal areas detected on the RGB optical data are visible on the corresponding LULC map. We now extract the littoral line and seafront line associated.

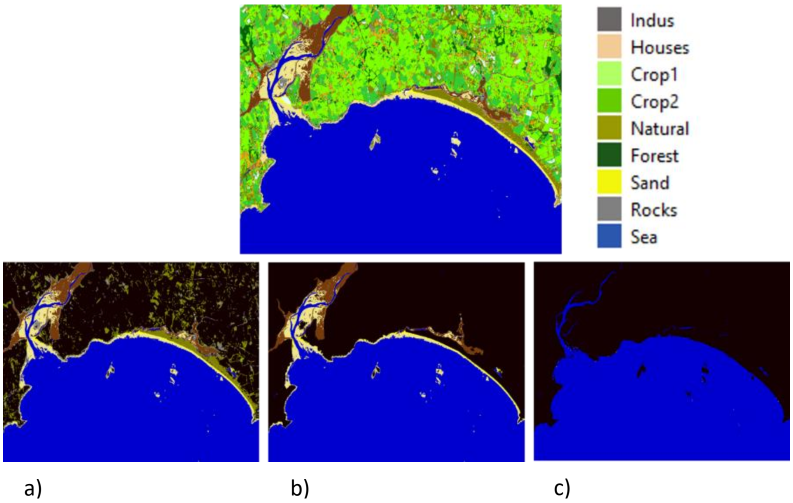


Figure 2.5: masks associated with the littoral line (a), seafront line (b) and land/sea limit (c) for Cork estuary.

3 Conclusion

End users are interested in shoreline indicators, these products include proxy-based (PSI) and datum-based (DSI) shoreline indicators. The littoral and seafront lines are produced for each location sites over the past 25 years frequency and production strongly relies on the erosion rate and land cover maps available for each study sites.

If both 1D products participate in the improvement of understanding of coastal morphodynamics for interannual evolution, our 2D classification maps proved to be great support in the comprehension of erosion context providing thanks to the temporal approach a representation of ocean dynamics such as tidal areas and sand bars dynamics. Therefore, the intermediate land classification product will also be available to the partners. LULC maps are needed to characterize the receptor for standard coastal risk management practices. Classes required for LULC map vary among end-user’s requirements.

Monitoring changes in land cover will allow to access any changes in the vulnerability of coastal areas and assess the coastal risk management of replacing hard structures by soft engineering and revegetation of the backshore or to inform and design regular maintenance and emergency works.

End of Document