ALES: a multi-mission adaptive sub-waveform retracker for coastal and open ocean altimetry

Marcello Passaro\textsuperscript{a}, Paolo Cipollini\textsuperscript{b}, Stefano Vignudelli\textsuperscript{c}, Graham Quartly\textsuperscript{d}, Helen Snaith\textsuperscript{e}

\textsuperscript{a}GSNOCS, University of Southampton, Waterfront Campus European Way, SO14 3ZH, Southampton, UK. Contacts: marcello.passaro@noc.soton.ac.uk, 00442380596488
\textsuperscript{b}National Oceanography Centre, Southampton
\textsuperscript{c}Consiglio Nazionale delle Ricerche, Istituto di Biofisica, Pisa
\textsuperscript{d}Plymouth Marine Laboratory
\textsuperscript{e}British Oceanographic Data Centre, Southampton

Abstract

Satellite altimetry has revolutionized our understanding of ocean dynamics thanks to frequent sampling and global coverage. Nevertheless, coastal data have been flagged as unreliable due to land and calm water interference in the altimeter and radiometer footprint and uncertainty in the modelling of high-frequency tidal and atmospheric forcing.

Our study addresses the first issue, i.e. altimeter footprint contamination, via re-tracking, presenting ALES, the Adaptive Leading Edge Subwaveform retracker. ALES is potentially applicable to all the pulse-limited altimetry missions and its aim is to retrack both open ocean and coastal data with the same accuracy using just one algorithm.

ALES selects part of each returned echo and models it with a classic "open ocean" Brown functional form, by means of least square estimation whose convergence is found through the Nelder-Mead nonlinear optimization technique. By avoiding echoes from bright targets along the trailing edge, it is capable of retrieving more coastal waveforms than the standard processing. By adapting the width of the estimation window according to the significant wave height, it aims at maintaining the accuracy of the standard processing in both the open ocean and the coastal strip.

This innovative retracker is validated against tide gauges in the Adriatic Sea and in the Greater Agulhas System for three different missions: Envisat, Jason-1 and Jason-2. Considerations of noise and biases provide a further verification of the strategy. The results show that ALES is able to provide more reliable 20-Hz data for all three missions in areas where even 1-Hz averages are flagged as unreliable in standard products. Application of the ALES retracker led to roughly a half of the analysed tracks showing a
marked improvement in correlation with the tide gauge records, with the rms difference being reduced by a factor of 1.5 for Jason-1 and Jason-2 and over 4 for Envisat in the Adriatic Sea (at the closest point to the tide gauge).

**Keywords:** Coastal Altimetry, retracking, subwaveform retracker, validation, tide gauge, satellite altimetry;

---

1. **Introduction**

Satellite altimetry, one of the most successful applications of remote sensing at the service of earth science and climate studies, is based on a simple radar principle. The fundamentals of satellite altimetry are comprehensively described in Chelton et al. (2001). A short pulse of radiation with known power is transmitted from a satellite towards the sea. The pulse interacts with the rough sea surface and part of the incident radiation within the altimetric footprint reflects back to the radar altimeter, which records the returned echo of the pulse. The power of the signal as received by the satellite is registered in a time series called a "waveform", sampled with a two-way travel time resolution of 3.125 ns; each resolution cell is known as a "gate". The acquisition depends on the functioning of the on-board tracker, which adjusts the altimeter observation window in time in order to keep the reflected signal coming from the Earth within the window. The output of the on-board tracker is the "Tracker Range", which is a first estimation of the distance between the satellite and the sea surface.

Each individual echo is strongly perturbed by Rayleigh noise (speckle) coming from the incoherent addition of signals from reflecting facets inside the satellite footprint (Quartly et al., 2001). Assuming that noise on consecutive waveforms is uncorrelated (which sets an upper boundary on the useful pulse repetition frequency of the altimeter (Walsh, 1982)), it can be reduced by averaging a number of waveforms prior to transmission (downlink) to the ground. Typical downlinked "high-rate" waveforms at ~ 20 Hz are an average of 100 consecutive samples at ~ 2 kHz (Quartly et al., 2001). In order to retrieve geophysical information from these waveforms, a processing step called "retracking" has to be performed on the ground (Chelton et al., 2001). The retracking process consists of fitting a model response to the real waveform. Over the open ocean, most waveforms are well described by the Brown-Hayne mathematical model (Brown, 1977; Hayne, 1980), which we recall in section 3 and which is the standard model used for retracking waveforms.
Waveforms that conform to the Brown-Hayne model present a fast-rising leading edge and a slowly-decaying trailing edge. Figure 1 shows the characteristics of an idealized open ocean waveform without noise, in order to define the parameters of interest for the retracking. The mid-point of the leading edge is chosen as an estimate of the distance between the satellite and the sea surface (range), since it is related to the sea surface mean position at nadir. The leading edge is therefore particularly important for the retracking, since it contains the parameters that a retracker usually estimates: the position of the mid-point of the leading edge with respect to the fixed nominal tracking point determined by the on-board tracker (Epoch $\tau$); the rise time of the leading edge, which is related to the significant wave height (SWH); the amplitude of the received signal, from which the backscatter coefficient $\sigma^0$ is derived and then related to the wind speed. The retracked range is computed by adding the Epoch to the Tracker Range.

The residual noise of real waveforms, particularly evident along the trailing edge, can influence the correct retrieval of the parameters of interest in the retracking process, since the waveforms deviate from the theoretical open ocean shape. This is known to happen in particular in the last 10 km from the coastline: at this distance, both coastal waters and raised land can give returns within the altimeter’s range window. In these cases, the waveform shape changes from that expected for a homogeneous surface. Several studies in recent years have classified coastal waveforms (Deng et al., 2002; Berry et al., 2010; Andersen et al., 2010). Two of the predominant coastal waveform classes are quasi-specular and multi-peak echoes (Idris & Deng, 2012). They are both connected with the presence of highly reflective targets (bright targets) whose signatures are seen tracing an hyperbola in consecutive waveforms (Gommenginger et al., 2011).

The physical features causing bright targets are still subject to debate. Tournadre (2007) states that signals coming from small targets like ships are only detected in the first gates before the leading edge, while land features such as islands can influence the entire waveform due to their high backscatter coefficient. In the same study, the fact that coastal bright targets are not present in every cycle is attributed to exposure of tidal flats by the tidal cycle. Recently, the same author has led a study demonstrating the detection of icebergs in the open ocean based on the presence of bright targets in the gates preceding the leading edge (Tournadre et al., 2012). Gomez-Enri et al. (2010) and
Scozzari et al. (2012) have investigated the bright targets over a coastal area near the island of Pianosa. They both concluded that the presence of patches of high reflectance is not a constant event and that they are most likely related to sea state and can be generated by unbroken wave crests and patches of calm water.

Figure 2 shows six along-track radargrams (stacks of consecutive hi-rate waveforms, one for each column) from Envisat track 416 flying over the northern end of the Adriatic Sea (map in figure 3). Land is shaded in grey. The colorbar codes the power of each gate for every waveform. Brighter features in the speckle noise along the trailing edge are seen progressing from later gates towards the leading edge and back, along hyperbolic paths as expected from simple geometrical arguments (Gomez-Enri et al., 2010).

Using geometry considerations described in Quartly (1998) and assuming that the actual tracking point is not significantly different from its nominal position, it is possible to compute the distance of the last gate of the waveform from the nominal tracking point, i.e. the radius of the satellite footprint at the end of the trailing edge, which is 7.3 km for Envisat and 8.3 km for Jason for calm seas. Depending on the elevation, land returns could still appear in the trailing edge, even if the surface is located outside the expected footprint, because their location could be equidistant with the ocean surface near nadir. Nevertheless, this would produce a more predictable hyperbolic feature than what is shown in the radargrams, where bright targets are not seen constantly at every cycle and their location and extent varies. It is therefore evident that the perturbations cannot be exclusively attributed to land in the satellite footprint.

Regardless of their origin, these features degrade the estimation of geophysical parameters based on Brown-Hayne retracking. To avoid this problem, different dedicated coastal retracking solutions have been suggested and tested in recent years. The latest strategies involve one or a combination of the following methods:

- the classification of the waveforms depending on their shape (Deng & Featherstone, 2006; Berry et al., 2010; Andersen et al., 2010; Yang et al., 2012)

- the use of empirical parameters and/or threshold values to model the waveform (Deng & Featherstone, 2006; Hwang et al., 2006; Bao et al., 2009; Lee et al., 2010)

- the adaptation of a different functional form for every kind of characteristic shape that the waveforms can assume (Berry et al., 1997; Andersen et al., 2010)
• the simultaneous processing of multiple waveforms to detect the bright targets prior to retracking (Gomez-Enri et al., 2010; Quartly, 2010)

• the addition of peaks to the Brown-Hayne functional form to model the presence of bright targets (Halimi et al., 2012)

• the retracking of subwaveforms, i.e. portion of waveforms not affected by the bright targets (Mercier et al., 2009; Guo et al., 2010; Yang et al., 2011, 2012; Idris & Deng, 2012)

Most of these attempts have been successful in demonstrating that meaningful geophysical estimates can be retrieved from data records which were previously disregarded. However, there are a number of still unresolved issues. The use of empirical and threshold retrackers lacks a physical description of the functional form; moreover, the choice of parameters might vary in different coastal conditions. The addition of further parameters (peaks) to the model can improve the overall fitting of the waveform without necessarily resulting in a better fitting of the three important parameters ($\tau$, SWH, $\sigma^0$), and also introduces additional complications in terms of numerical convergence of the estimation. The use of different retrackers depending on conditions and the switching from one to another, requires an accurate quantification of biases amongst them, to avoid "jumps" in the retrieved parameters. Even the simple distinction between open and coastal ocean to apply different retracking strategies raises questions about where the boundary should be set, and what are the consequences of introducing such a discontinuity.

In this study we present the Adaptive Leading Edge Sub-waveform (ALES) retracker, which will overcome the above difficulties as it can be applied both over open and coastal ocean without discontinuity. ALES represents an evolution of the family of the subwaveform retrackers, in that it adapts the width of the sub-waveform according to the SWH. ALES is designed in order to maintain the same degree of precision both in open ocean and along the coasts. As we will demonstrate, ALES performs only marginally worse than the Brown-Hayne retracker over waveforms that comply with the Brown-Hayne model, but it outperforms the Brown-Hayne retracker for the vast majority of open ocean waveforms affected by bright targets, as well as in coastal waters. One of ALES strengths lies in the fact that it does not involve any change in the retracking method, hereby avoiding any source of potential internal bias. It succeeds in retracking
waveforms where the trailing edge is contaminated, but a distinct leading edge is still observable.

The rationale for ALES comes from recent studies on leading edge retrackers, but also from concepts deeply embedded in the design of the first altimetric missions. The tracker system designs of the SEASAT, GEOSAT and TOPEX missions, while using the rising leading edge as the basis of precise height measurement, estimated the height error (degradation of measurement precision with increasing wave height) using a number of gates increasing with increasing wave height (Marth et al., 1993). The variation in width of the fitting window with SWH makes ALES widely applicable to different environmental conditions as we demonstrate over two regions with very different sea state. Moreover ALES is applicable to waveforms from different altimeters, as we show in this study with data from Envisat, Jason-1 and Jason-2: to our knowledge this is the first time that the same coastal-dedicated retracker has been applied and validated for three different altimetry missions.

Section 2 presents the datasets that have been used in the study. Section 3 describes the mathematical details and the functioning of ALES. The sea level estimation from ALES is then validated for the three altimetric missions in section 4, where the results are presented and discussed. Section 5 draws the conclusions and the outlook for future research enabled by ALES.

2. Dataset

Two areas of study (shown in 3) were selected to verify and validate ALES: the Gulf of Trieste, in the North-West Adriatic Sea, and Mossel Bay, along the Indian Ocean coast of South Africa. The areas are representative of two different morphological and oceanic conditions.

The Gulf of Trieste is a small shallow bay with average depth: 17 m, maximum depth: 25 m and width: 20 km. The maximum tidal amplitude is 81 cm (Querin et al., 2006). The cyclonic gyre (anticlockwise) that characterises the main circulation is induced by currents flowing northwards along the Istrian coast and is modulated daily by the local wind field (Querin et al., 2006). The rest of the Northern Adriatic basin is also a shallow sea, less than 100 m deep. It is characterised by a weak cyclonic circulation that reaches peaks of 25 cm/s along the Italian coast of its western boundary (Poulain, 2001).
Wind strength and direction is particularly important because of its influence on the sea surface roughness. The most important wind in the Gulf of Trieste in terms of frequency and intensity is called the Bora (from East-North-East), with further contributions to atmospheric circulation from Scirocco (direction South-East to South-South-West) and breezes from land and sea. The Bora is a strong katabatic wind which can reach speeds of over 10 m/s, especially during autumn/winter and has an annual mean velocity of 6 m/s (compared to 2.1 m/s for the other winds in the region) (Stravisi, 1977).

Mossel Bay is a much more open 25 km wide gulf, affected by the large circulation features of the South African coast. From the coast, the Agulhas Bank extends for roughly 200 km with depths that do not exceed 200 m and near-surface currents in the range of 0-0.5 m/s. Then the continental slope is found, where the prevailing westward warm Agulhas Current reaches velocities of 2.50 m/s and the depth rapidly reaches over 4000 m (Boyd & Shillington, 1994). The amplitude of spring tide reaches 2.40 m (South African Navy, 2007). Wind forcing is particularly important along the coast and influences the surface currents of the bank. Easterly winds prevail in the summer, causing coastal upwelling, while westerly winds predominate in winter (Schumann, 1992).

In these two locations, waveforms from 5 tracks from 3 different altimetry satellites were retracted with ALES. The waveforms are provided as part of the SGDR (Sensor Geophysical Data Record) product for each mission, which also includes the level 2 geophysical estimates of height, SWH and wind based on the Brown-Hayne retracker, together with atmospheric and geophysical corrections to be applied to the data. They are distributed at 18 Hz rate for Envisat and 20 Hz rate for Jason-1 and Jason-2, representing a measurement every 300-350 metres.

As a reference and comparison, data from CTOH (Centre for Topographic studies of the Ocean and Hydrosphere, http://ctoh.legos.obs-mip.fr/altimetry) have been downloaded. In particular, the 1-Hz (a measurement approximately every 7 km) along track altimetry output from X-TRACK processor is used. At the time of this research, data were available for all the tracks except for Envisat pass 687. Data were developed, validated and distributed by the CTOH/LEGOS, France. The X-TRACK processor does not retrack the waveforms; it is a post-processing technique aimed at improving quality and quantity of coastal altimetry measurements through improved tidal and atmosphere forcing correction, data editing and filtering. A detailed description can be found in Roblou
et al. (2007) and Roblou et al. (2011). X-TRACK data are widely used and applications already include the Mediterranean Sea (Birol et al., 2010; Bouffard et al., 2011; García et al., 2012), the Spanish coast (Le Hénaff et al., 2011; Dussurget et al., 2011; Herbert et al., 2011), the West Florida Shelf (Liu et al., 2012), Drake Passage (Ferrari et al., 2013) and the Solomon Sea (Melet et al., 2010).

To create a time series, data points along the satellite tracks have to be collinear: it is necessary to have measurements at the same geographical location for each cycle. Nominal tracks were therefore created for this study by taking as a reference the CTOH 1-Hz tracks, neglecting the across-track displacement of different passes along the same track, which is normally less than 1 km. Each interval between consecutive 1-Hz data points was divided in order to obtain 20 equidistant nominal locations, along which the altimeter data from SGDR and ALES for each cycle were then linearly interpolated.

Two tide gauges were used for validation of ALES output. They are located at Trieste (PSMSL Station ID 154, Latitude 45.647361° N, Longitude 13.758472° E) and Mossel Bay (PSMSL Station ID 910, Latitude -34.178611° N, Longitude 22.135278° E).

The Trieste tide gauge belongs to the Italian national tide gauge network operated by ISPRA (High Institute for Environmental Protection and Research). The station is equipped with a gauge that records sea level referred to the zero level of the Italian altimetric network. The Italian altimetric network, created by IGM (Istituto Geografico Militare) is the vertical reference (datum) to which all heights are measured around Italy. The zero of the Italian altimetric network is the mean sea level, measured in Genoa during the years 1937-1946 (Gamboni, 1965). For the Trieste station, it was possible to refer the sea level measurements to the WGS84 ellipsoid, since the Italian zero level at the tide gauge is estimated to be 44.760 m above it (Stravisi & Purga, 2005; Becker et al., 2002). The sampling frequency is adjusted to provide data every 10 minutes and/or hourly.

Hourly data for Mossel Bay are distributed by the GLOSS/CLIVAR (Global Sea Level Observing System / Climate Variability and Predictability) fast sea level data service. Since we have been unable to find sources referring those measurements to the ellipsoid, they have been used for relative calibration only.

Figure 3 shows the area of study, including the satellite tracks and tide gauges locations. It is relevant to point out the flight direction of each satellite as it flies over each region, since land-to-sea and sea-to-land transitions might influence the behaviour
of the on-board tracker in different ways. In the Adriatic, Envisat pass 416 (Env 416) is a descending (North to South) track, as is Jason-2 pass 196 (J-2 196), while Jason-1 pass 161 in its interleaved orbit (J-1 161) is ascending (South to North). All three tracks cross part of the Gulf of Trieste and fly over the Istrian peninsula. For J-1 161 and Env 416, the stretch of land in Istria accounts for a few km along-track, while J-2 196 flies over the full extent of the peninsula. J-2 196 is particularly challenging because it also flies over Marano Lagoon (classified as inland water), which is less than 3 m deep, and then crosses a 0.5 km wide sandbar before entering the Adriatic Sea (Ferrarin et al., 2009).

Of the two tracks we have used along the coast of South Africa, Envisat pass 687 (Env 687) is an ascending track, while Jason-2 pass 198 (J-2 198) is descending. J-2 198 flies over the same bay where the tide gauge is located, while Env 687 passes over the other side of Mossel Bay peninsula.

Table 1 summarises the characteristics of the dataset and in particular the number of waveforms that were missing from the along-track data and the number of waveforms for which no altimetry parameter estimation was available in the SGDR. The number of missing waveforms is obtained by checking how many times there are no waveforms that fall in the interval between each latitude-longitude point of the nominal tracks. Missing waveforms are more frequent in the Jason missions due to failures of the on-board tracker, an event that particularly occurs near the transition from land to sea (Brooks et al., 1998), which explains why the loss is concentrated in the first 10 km from the coast for the two descending tracks of Jason-2, with up to 70% of data missing for the latitude-longitude points located closest to the coast.

The Envisat altimeter was designed to operate at different bandwidths over different surfaces. The bandwidth preferred for the ocean was 320 MHz, in order to guarantee the highest possible resolution. Nevertheless in the transition zone it can happen that returned echoes are acquired using a different bandwidth. For Envisat data, only SGDR output when operated at 320 MHz bandwidth are considered valid. ALES only retracks these waveforms.

A detailed list of all the cycles reprocessed for each satellite track and the related time interval is provided in table 2.
3. Retracker

3.1. Functional form

ALES adopts the Brown-Hayne theoretical ocean model (Brown, 1977; Hayne, 1980), the same as in use for the open-ocean retrackers, which describes the average return power of a rough scattering surface (i.e. what we simply call waveform). The return power $V_m$ is

$$V_m(t) = a_\xi P_u \frac{[1 + \text{erf}(u)]}{2} \exp(-v) + T_n \quad (1)$$

where

$$a_\xi = \exp\left(-\frac{4 \sin^2 \xi}{\gamma}\right) \quad \gamma = \sin^2(\theta_0) \frac{1}{2 \cdot \ln(2)}$$

$$\text{erf}(x) = 2 \frac{1}{\sqrt{\pi}} \int_{0}^{x} e^{-t^2} dt$$

$$u = \frac{t - \tau - c_\xi \sigma_c^2}{\sqrt{2} \sigma_c} \quad v = c_\xi \left(t - \tau - \frac{1}{2} c_\xi \sigma_c^2\right)$$

$$c_\xi = b_\xi a \quad b_\xi = \cos(2\xi) - \frac{\sin^2(2\xi)}{\gamma} \quad a = \frac{4c}{\gamma h \left(1 + \frac{h}{R_e}\right)}$$

$$\sigma_c^2 = \sigma_p^2 + \sigma_s^2 \quad \sigma_s = \frac{\text{SWH}}{2c}$$

where $c$ is the speed of light, $h$ the satellite altitude, $R_e$ the Earth radius, $\xi$ the off-nadir mispointing angle, $\theta_0$ the antenna beam width, $\tau$ the Epoch with respect to the nominal tracking reference point, $\sigma_c$ the rise time of the leading edge (depending on a term $\sigma_s$ linked to $\text{SWH}$ and on the width of the radar point target response $\sigma_p$), $P_u$ the amplitude of the signal and $T_n$ the thermal noise level.

In practice, the model in equation 1 is a raised sigmoid $\frac{[1 + \text{erf}(u)]}{2}$ describing the increasing power in the waveform leading edge and the subsequent plateau, multiplied by a
negative exponential which models the reduction of power in the waveform tail (decay), plus thermal (additive) noise $T_n$. The amplitude of the signal $P_u$ is attenuated by a term $a_\xi$ dependant on mispointing $\xi$. $P_u$ can be converted into a measurement of the backscatter coefficient $\sigma^0$ on the basis of the instrument calibration. Note that the significant wave height $SWH$, in addition to affecting the rise time of the waveform, also has a small effect on the sigmoid location (variable $u$) and on the waveform decay (variable $v$), via the term $c_\xi \sigma^2$.

A complete physical description of the functional form can be found in Brown (1977) and Gommenginger et al. (2011).

3.2. Estimator

The model in 3.1. is fitted to the real signal in order to estimate the parameters of interest. Chelton et al. (2001) suggested the use of least squares rather than maximum likelihood estimation (MLE), because the latter assigns too much weight to portions of the waveform with low signal power, and is therefore unreliable due to unmodeled errors in the waveform samples. For Envisat, Jason-1 and Jason-2 the optimal parameters are found at each iteration by developing the total cost function in a Taylor series at first order, which involves the computation of weighted partial derivatives; the method is also known as minimum mean square estimator (MMSE) (Gommenginger et al., 2011; Zanife et al., 2003).

ALES adopts an unweighted least-square estimator whose convergence is sought through the Nelder-Mead (NM) algorithm. NM is a simplex optimization method that does not use the derivatives of its cost function, whilst it searches for the minimum in a many-dimensional space. Specifically, considering $m$ parameters to be estimated, given that a simplex of dimension $m$ is a polytope of the same dimension and with $m+1$ vertices characterised by $m+1$ cost function values, NM generates at each step a new point whose cost function is compared with its value at the vertices. If it is smaller, the point becomes a vertex of the new simplex and a new iteration is generated (Nelder & Mead, 1965). Convergence is reached when the diameter of the simplex is smaller than a specified tolerance, set at $1 \times 10^{(-10)}$ in this study. For ALES, 600 is the maximum number (Nmax) of iterations allowed to reach the tolerance. As already demonstrated in Halimi et al. (2012), the method is also very effective in terms of speed: the average CPU time needed to process each oceanic waveform with the MATLAB code generated for ALES is
of the order of 0.1 seconds; the drawback of avoiding a waveform classification is that it can take up to few seconds for each waveform where no clear leading edge is detectable, such as on land.

Previous studies such as Olsson & Nelson (1975) have shown the properties of NM, and Halimi et al. (2012) examined the root mean square error in the parameter estimation for Brown retracking and found better performances compared with the Newton-Raphson strategy, which is also an iterative method based on the derivatives such as MMSE. The downside is that NM does not provide the Fisher information matrix to assess the confidence interval of the fitted parameters, nevertheless an evaluation can be performed by comparing the reconstructed waveform with the original returned echo (Halimi et al., 2012).

3.3. ALES strategy

Initially, each waveform is normalised by the maximum value of an 8-point average computed at each gate. Then, ALES performs the fitting in two passes; first it detects the leading edge of the waveform, as explained in 3.3.1 below, and computes a first estimate of $\tau$, SWH and $\sigma_0$ by fitting the model in 3.1 to the leading-edge; the SWH value is used as explained in 3.3.2 to select a SWH-dependent subwaveform window on which to perform a more precise estimate of the three parameters in the second pass of the algorithm. A flow diagram of the main steps followed by ALES to retrack each waveform is shown in figure 4. For brevity we will refer to the start gate number and end gate number of the sub-waveform window as the startgate and stopgate, respectively. A subwaveform will therefore be defined as the part of the waveforms whose first and last gates are identified by startgate and stopgate.

3.3.1. First pass: leading edge detection and fitting

The first estimation is performed on a window that includes the leading edge of the echo, entirely excluding the trailing edge.

Thermal noise is estimated from an average of the first few gates and removed from the signal. For Envisat, gates 5 to 10 are chosen, since the first gates are affected by aliasing due to the application of a digital filter that wraps around the high frequency components of the received pulse into the low frequency waveform samples. In Jason-1 and Jason-2, the first and the last gates are already excluded from the waveforms (104
gates are given, rather than the full 128 as in Envisat), therefore we estimate thermal
noise from the average of gates 1 to 5. For the same reason, ALES startgate is gate 5 for
Envisat and gate 1 for Jason-1 and Jason-2.

The main leading edge detection is based on the difference between consecutive gates.
The algorithm searches for the beginning of the leading edge (edge foot): this is taken
to be the first gate where the difference of two consecutive gates is positive and greater
than 1% of the normalization factor, i.e. when:

\[ Dwf > 0.01 \]

where \( Dwf \) is the vector containing the differences between consecutive gates and
numbers are expressed in normalised power units. The 1% threshold is arbitrary and it
is defined to detect the reception of a return echo.

Then the algorithm finds the end of the main leading edge (edge top) at the first
following gate where the difference of consecutive waveform values becomes negative, i.e.
when:

\[ Dwf < 0 \]

The stopgate for this first pass is taken as the edge top gate plus one. Due to noise,
the exact location of the end of a leading edge is not precisely defined in a waveform.
This can lead to a problem of accuracy in the case of a simple leading-edge retracker.
However, the stopgate is here only used to define the subwaveform for the first estimation
of Epoch and SWH, whose final values are the result of the second estimation, which is
independent of the stopgate.

Returns from discrete raised surfaces, such as the decks of ships, can sometimes be seen
as small power spikes in the gates preceding the leading edge. These can be erroneously
detected as leading edges according to criterion 2. To avoid such erroneous detections,
edges whose power drops below 10% of the normalisation factor within 4 gates from the
edge top are excluded.

The subwaveform is then fitted as in section 3.2. When convergence is not reached in
\( N_{\text{max}} \) iterations, the estimation window is widened by increasing the stopgate by 1 (i.e.
widening the window towards the waveform tail), and so on, until convergence is reached.
Epoch and SWH from the leading edge sub-waveform estimation are then used to select a wider sub-waveform for the second pass.

3.3.2. Second pass: selection of sub-waveform and precise final estimation

For best accuracy the subwaveform width for the second pass must be optimised such that it fully includes all gates comprising the leading edge, but with minimal contribution from the trailing edge, where artefacts such as bright target responses may prevent equation 1 from accurately describing the shape. In effect the issue is one of defining an appropriate stopgate for the second pass retracking based upon the SWH estimates from the first pass. The relationship between SWH and stopgate was derived from Monte Carlo simulations. For each value of SWH ranging from 0.5 to 10 m in steps of 0.5 m, 500 high-rate waveforms were simulated with the model in 3.1 adding realistic Rayleigh noise. The resulting waveforms were retracked using the classic Brown-Hayne model in 3.1 over the entire waveform, and then over sub-waveform windows with startgate=1 and variable stopgate, and the RMS errors (RMSE) computed.

The difference of the RMSEs between the "full waveform" estimate and the sub-waveform estimates is displayed in figure 5 as a function of the stopgate position. Upper left, upper right and lower left panels show the results for Epoch, SWH and amplitude estimation. The x axis is, in practice, the width of the sub-waveform, expressed as number of gates from the tracking point to the stopgate. The results for each SWH level are coded in different colours. For all three parameters, the curves converge asymptotically to the full waveform estimates, as expected for this idealised case of "pure-Brown" response of the ocean surface. The amplitude is the parameter that needs the smallest amount of gates to be properly estimated - and indeed in some cases its estimate can be marginally more accurate by discarding part of the tails, as indicated by the positive difference of RMSEs. Epoch and SWH estimations show similar patterns, although the curves are less smooth for the SWH. This is also expected, since SWH retrieval is normally noisier in the retracking of real waveforms. In the lower right panel, a linear relationship is derived between SWH and the number of gates needed using a 1 cm RMSE tolerance compared to the theoretical precision of a "full waveform" estimation. The choice of the 1 cm tolerance is a trade off between retaining an "open-ocean" level of precision and avoiding perturbations in the trailing edge. We consider 1 cm a good compromise since it is much less than the estimated noise on high rate range measurements found in previous studies.
on Jason-1, Jason-2 and Envisat (Faugere et al., 2006; Thibaut et al., 2010).

The parameters $h$, $\theta_0$ and $\sigma_p$ (defined in section 3.1) change depending on the mission. Therefore the same Monte Carlo simulation was run for both Envisat and Jason, resulting in slightly different coefficients of the linear relationship derived (figure 4 refers to Envisat). From these results, the stopgate of the ALES estimation window is calculated using the following equations:

$$\text{Maximum gate} = \text{Ceiling}(\text{Tracking point} + 2.4263 + 4.1759 \times \text{SWH})$$  \hspace{1cm} (4)

for Envisat and

$$\text{Maximum gate} = \text{Ceiling}(\text{Tracking point} + 1.3737 + 4.5098 \times \text{SWH})$$  \hspace{1cm} (5)

where tracking point and SWH are from the first pass. The final estimation of the three parameters $\tau$, SWH and $\sigma^0$ is made from this window width at the second pass.

Examples of three typical fittings of real waveforms (blue line) by ALES (dashed red line) are provided in figure 6 for open ocean with low SWH, coastal ocean with corrupted trailing edge and open ocean with high SWH.

3.3.3. Mispointing angle

Every radar altimeter has a degree of inaccuracy in pointing perpendicular to the ocean surface. The slope of the trailing edge in the theoretical Brown functional form depends on the mispointing angle, which measures how much the antenna pointing deviates from nadir. The most significant effect is in the estimation of the signal amplitude, reducing the apparent backscatter coefficient (Wingham, 1988; Quartly, 2009).

In its current version, ALES does not estimate mispointing. Values are taken from the SGDR products. In standard three-parameter retracking, such as for Envisat, the angle is computed from the slope of the logarithm of the trailing edge (Amarouche et al., 2004). In the current version of Jason-1 and Jason-2 SGDR, the mispointing is estimated in the retracking process (from the MLE4 retracker) together with Epoch, SWH and amplitude. We use this output as input for our three-parameter model, and this should not influence the correct range retrieval for angles smaller than $0.3^\circ$ (Dorandeu et al., 2004). Where no
estimate from the SGDR is available, for example where SGDR retracking failed, we take
the last valid measurement. The mispointing values are then smoothed over 3 seconds.
These are reasonable approximations considering that Amarouche et al. (2004) suggested
that a smoothing over 30 seconds should be applied when mispointing is calculated from
the waveform. It is therefore appropriate to always smooth the value for the mispointing
to avoid features in the trailing edge affecting ALES performance through the mispointing
values used.

4. Validation and discussion

This section presents the validation of ALES range retrieval. To check the data
quality, time series of sea level variability measured from altimetry and tide gauges can
be compared by analysing correlation and root mean square (RMS) (Fenoglio-Marc et al.,
2010). The variable that we extracted for the comparison is the Total Water Level
Envelope (TWLE), defined as:

\[
TWLE = \text{Orbit altitude} - \text{Corrected Range} - \text{Mean Sea Surface} - (\text{Solid Earth Tide} + \text{Load Tide}) \quad (6)
\]

where

\[
\text{Corrected Range} = \text{Range} + \text{Instrumental corrections} + \text{Dry tropospheric correction} +
\text{Wet Tropospheric Correction} + \text{Sea State Bias} + \text{Ionospheric correction} \quad (7)
\]

TWLE represents the combined effect of ocean tides and atmospheric forcing in add-
tion to the sea level anomaly with respect to the mean sea surface. It is therefore
particularly useful for applications that need an estimate of the total sea level above the
mean, such as the analysis of storm surges. The DTU10 mean sea surface (Andersen,
2010) was subtracted from the retrieved sea surface height for all the tracks. In order
to compare the altimeter data with the TGs, the effect of atmospheric pressure on the
sea level was not applied to either dataset (Fenoglio-Marc et al., 2004). While Envisat
and Jason-1 SGDR provide estimations of the range from a single oceanic retracker, the
latest version of Jason-2 gives two estimations for each waveform, coming from MLE3
and MLE4 retrackers (Thibaut et al., 2010). To be consistent with the comparison with
Jason-1, only the MLE4 estimates for Jason-2 are used in this study.
A numerical summary of the comparative analysis with the tide gauges for each track is presented in table 2. The outliers are defined according to section 4.2. For each track, the median value of the along-track correlation is shown. For J-2 198 and Env 687 in Mossel Bay (figure 3), since there was no independent estimate of TG height relative to the ellipsoid, RMS values correspond to the median value of the along-track RMS of the difference between TWLE and TG sea level height anomaly (with the mean of the difference removed), while for J-2 196, J-1 161 and Env 416 in the Adriatic Sea we report the RMS of the difference of the absolute sea level heights above the ellipsoid at the closest point between TG and satellite tracks. The altimetric absolute sea level heights are obtained by adding the mean sea surface to the TWLE defined in equation 6.

4.1. Validation of raw data

The first validation test was performed on all the available along-track range estimations from both SGDR and ALES. The aim was to determine for each latitude longitude location the maximum number of cycles of data that could be retained whilst guaranteeing a correlation with the TG time series of at least 0.9. The test was performed in an iterative way: for each location, the correlation of the TG time series with the entire set of altimetry retrievals was checked; if the correlation coefficient was lower than 0.9, then the cycle with the maximum discrepancy between TG value and altimeter retrieval was excluded. The output assesses the general quality of the retracked altimetry values against the available SGDR product. It is important to perform the analysis before any outliers exclusion, in order to quantitatively evaluate whether ALES, compared to standard products, can provide a higher number of accurate estimations at each location. The same analysis was performed for CTOH 1-Hz locations.

Figure 7 and 8 show the results for the tracks in the Adriatic and the tracks near South Africa. Land areas are shaded in grey. The x-axis represents the along-track progression of each track in latitude. The latitude of the tide gauge is specified by a black dot. Red squares characterise the ALES dataset, while SGDR products are in blue squares and CTOH points are seen as cyan dots. The distance from the closest coastline is specified by a green line which refers to the y-axis on the right.

Env 416 height retrieval is particularly problematic in the gulf area, between 45.5 and 45.8° N. Here ALES shows improvements particularly in the sea-to-land transition, with over 60% of cycles highly correlated in locations closer than 2.5 km to the coast.
where the amount of highly correlated cycles in the SGDR product decrease significantly.

After roughly 3.7 km from the stretch of land in Istria, both ALES and SGDR have more than 80% of cycles highly correlated, but along the rest of the track ALES constantly maintains a higher percentage.

The improvement of data quality in the gulf area is even more evident considering J-1 161. The percentage of ALES highly correlated cycles is constantly well above 60%, a threshold that is hardly overtaken by SGDR in the same locations. Before the Istrian stretch, availability of good ALES data is constantly over 80% of the cycles until 2.5 km from the coast, while the percentage for the SGDR drops below 80% by 7 km from the Istrian peninsula. For both J-1 161 and Env 416 in the Gulf of Trieste, CTOH had no data for any cycle. The first available measurements are situated more than 10 km along-track from the Istrian peninsula and several cycles contained empty data, as evident by the low percentage in the figure.

Along the descending J-2 196 track, ALES retrieves more than 80% of highly correlated cycles within 7 km along-track from the islet that marks the beginning of the lagoon (visible on the plot as a dashed black vertical line), while the SGDR data take 13 km (along-track distance) to get to a comparable level. It is particularly striking how ALES is also able to capture more valid data also inside the lagoon, showing that ALES could potentially be used for inland water sea level measurements.

In the tracks that pass over South Africa, which fly over a wide open ocean area, the results for SGDR and ALES are extremely similar. Nevertheless the performances of SGDR slightly decrease at 11 km from the coast, while ALES points are highly correlated to within 1.5 km. Along J-2 198, the parallel degradation of the two datasets in proximity to the coast is explained by failure of the Jason on-board tracker, an event that particularly occurs in the transition from land to sea (Brooks et al., 1998).

4.2. Outlier detection

In order to further analyse the correlation and the RMS between in-situ sea level observations and ALES estimations, outliers have been detected and excluded. Validity does not imply accuracy and therefore the number of valid points does not correspond to the total number of "correct" estimates; nevertheless the outlier detection is applied to exclude extreme values of ALES output which would alter the correlation and RMS analysis performed on the full time series.
For the sake of comparison, the same flagging criteria have been used for both SGDR and ALES data. It is preferable to use separate quality control for each of the three retracked parameters (TWLE, SWH and $\sigma^0$) in order to exclude as many poor estimations as possible. Although a flag on TWLE is not suggested in the SGDR and ocean tides are not removed from this quantity, tidal signal is available from the tide gauges. Therefore a specific flag has been derived for this research taking into account the suggested flag for sea level anomaly in the Envisat product. SWH and $\sigma^0$ suggested flagging criteria are the same as the ones used for both Envisat and Jason products (Picot et al., 2003; Soussi & Femenias, 2006). Table 4 lists and describes the adopted flagging criteria used for outlier detection. For Envisat, since the SGDR product does not provide the high-rate values of SWH and $\sigma^0$, only the flag on TWLE could be applied.

Figures 9 and 10 show the along-track distribution of outliers in the two considered regions. The most problematic area is the Gulf of Trieste and the reason lies in the fact that a large part of these tracks is closer than 10 km from the coastline (with multiple transitions from land to sea and sea to land), where the majority of Jason SGDR estimations are outliers. Envisat has the smallest number of outliers among the different missions. It must also be noticed that SGDR in the Adriatic has more valid points than ALES. Considering the low correlation of SGDR (see section 4.3) along this track in the locations where the outliers are less than for ALES, we conclude that some of the SGDR poor estimations are not extreme enough to be eliminated by the outliers detection.

Since in Jason missions the outlier detection can be extended to SWH and $\sigma^0$, we assume this comparison to be more meaningful. In J-1 161 ALES succeeds in keeping the outliers below 25% even very close to the coastline, where SGDR exceeds 60%. In J-2 196 the number of ALES outliers is almost always less than the number of SGDR outliers, and between 45.6 and 45.7 °N there are less than 20% outliers, whereas SGDR reaches over 60%. In J-2 198 the percentage of outliers rises in the last 6 km from the coastline, at the same pace both in SGDR and in ALES. Here waveforms are highly corrupted and a leading edge is difficult to recognise, as a result of the problematic land-to-sea transition of the Jason altimeter.

It has already been said that erroneous estimations are due to waveforms that do not conform to the Brown model, among which those with a moving bright target in the
trailing edge are common in coastal areas, as showed in previous studies such as Halimi et al. (2012), where waveforms are classified according to their shape. We argue that the higher percentage of outliers in the Gulf of Trieste compared to Mossel Bay is connected with the shape of the coastline and its influence on the sea state. Depending on wind intensity and direction in the Gulf of Trieste the rugged coastline can create patches of sheltered calm water visible as bright targets in the radargrams, such as the ones of figure 2. Further explanation must be found in the analysis of the wind direction, frequency and intensity, but also of the wave height and in the interaction between wind and currents, since all these factors affect the sea surface roughness (Lange et al., 2004). Moreover, film slicks on the sea surface are common in the coastal ocean and are mainly related to surfactant accumulation and upwelling fronts (Ermakov et al., 1992; Ryan et al., 2010). They can extend for many kilometres in calm conditions and create patches of very low surface roughness that can be seen as bright targets in the radar signal. Oil slicks from ships are also to be investigated as a source of highly reflective water: both North Adriatic and South African coasts have high shipping activity, but the slow surface currents of the Adriatic and in particular of the Gulf of Trieste might enhance the permanence of the slicks in that area. Phytoplankton blooms can also have a strong effect, but their signal would be highly seasonal and more sporadic than the observed recurrence of bright targets.

4.3. Correlation coefficient

Once the outliers were removed, the correlation between ALES TWLE and TG was computed. For comparison, the correlation coefficient (r) was also computed for SGDR output from the same waveforms. As table 1 shows, SGDR data display missing values in locations where waveforms are present. For each location, in order to produce an unbiased comparison, we only consider correlation when both SGDR and ALES estimates are available.

Along-track spatial variation of r for the different tracks is displayed in figures 11 and 12. ALES improvement in correlation compared to SGDR is widespread. In the open seas, r for ALES is constantly over 0.8, with a few drops that may be due to undetected erroneous TWLE estimations, but may also be due to imprecise corrections when the drops are seen also in r for SGDR. In the Gulf of Trieste, r for ALES is always higher than for SGDR for distances greater r than 2.5 km from the coastline, and it is always
above 0.6 for J-1 161 and J-2 196. The low values of r for Env 416 suggest that a more
careful quality control of the estimations will be necessary in a post-processing phase; for
example, separate quality control of the SWH and \(\sigma^0\) estimations, which has not been
done in this study for Envisat due to the lack of 20Hz estimations in the SGDR product,
as explained in the previous section.

Env 687 in South Africa presents a steady high correlation of over 0.9 until within 2.5
km of coastline, but the SGDR data have a significant drop between 10 km and 5 km
from the coastline, a degradation that is not seen in ALES data. Along J-2 198 SGDR
data also have a drop at 10 km from the coastline, while ALES maintains r above 0.8.
ALES presents a single low r value in the open ocean, a closer investigation of this case
revealed the presence of a false leading edge, indicating a future area for improvement in
the routine described in section 3.3.1.

The comparison with CTOH data is extremely promising: ALES data have the same
level of correlation with the tide gauges despite being a high-rate non-filtered product.
A higher correlation of the tide gauges with CTOH 1-Hz data was expected as these
data have already been post-processed and therefore outliers have been excluded in the
calculation of 1-Hz mean from the 20 Hz measurements. Reaching the same correlation
as CTOH values without any filtering or selection means that high rate values with
appropriate flagging can be effectively used to produce altimetric data at finer resolutions
and in addition areas where no CTOH output is available, such as the innermost part of
the Gulf of Trieste, can still be described by ALES with unprecedented accuracy.

4.4. Root Mean Square Difference

The absolute validation against coastal tide gauges in terms of sea level can help to
assess whether the comparability between in-situ data and altimetry is increased by the
use of improved retracking strategies. Figure 12 shows the RMS of the difference (RMSD)
between the time series of the absolute sea level height above the ellipsoid WGS84 (major
radius \(a=6378137\) m, flattening \(f=298.257223563\)) for the tracks in the North Adriatic
and the equivalent time series at Trieste TG. Since the reference ellipsoid of the satel-
lite altitude in Jason and DTU10 mean sea surface is TOPEX ellipsoid (major radius
\(a=6378136.3\) m, flattening \(f=298.257\)), values are corrected for the difference between
the two ellipsoid models in order to refer each parameter to WGS84. The heights from
satellite altimetry are also corrected for the absolute biases found in Bonnefond et al.
Minimum RMSD values are expected to be found using the along-track point nearest to the TG, although this depends on the slope of the local geoid and local currents. Heights measured by the tide gauges and by the altimeters are always different since the altimeters do not fly exactly over the tide gauge and the tide gauge is a point measurement, while a radar altimeter considers the entire footprint. For a more precise estimation, altimetry calibration processes take into account the difference in geoid height between the track and the TG, but this is done only at dedicated sites where regional high-resolution geoid models are available (Bonnefond et al., 2011). Moreover even presently available high-resolution global geoid models (such as EGM08) are not suitable for this purpose in proximity of the coast (within 20 km) (Bonnefond et al., 2012). In our comparative analysis we do not aim to calibrate a sensor and we therefore do not apply a correction for geoid gradient. However, to help the evaluation, figure 13 also includes the difference between the DTU10 mean sea surface at the TG and the along-track.

In the open Adriatic, ALES RMSD values decrease steadily towards the tide gauge for Env 416 and for J-1 161, with worsening performances for Env 416 in the last 2.5 km. In the gulf, RMSD values are more variable, in particular for Env 416, but the comparison with the SGDR output for the same waveforms highlights the improvements of our new retracking system. For J-2 196, the absolute RMSD is more difficult to evaluate, because of the greater distance (over 30 km at the closest point) from the tide gauge. ALES RMSD accurately follows this shape from 4 km of distance from the Istrian coast until the proximity of the lagoon, in locations where SGDR values are extremely different from the tide gauge measurements.

Since the ellipsoidal height of Mossel Bay TG was not available, a relative analysis was performed by computing the RMS of the difference between TWLE and tide gauge values. The results are shown in figure 14 and are very comparable to the correlation analysis.

4.5. Distribution and bias analysis

One of the main aims of this study was to assess the validity of the same retracker for both open ocean and coastal data. It is important therefore that the high rate estimations of the new retracker are not noisier than the standard SGDR product. We consider the differences between consecutive high-rate TWLE values as a good estimation of noise,
since TWLE is not supposed to change significantly in 300 to 350 metres, which is the
distance between one measurement and the next.

TWLE consecutive differences were computed for ALES, once the outliers had been
removed. For comparison, the same calculation was performed for SGDR for the same
waveforms. The histograms of consecutive absolute differences of TWLE are shown for
each track in figure 15. Table 4 describes the characteristics of the distributions in terms
of mean, standard deviation and percentiles (1st and 99th of the whole probability density
function).

The majority of consecutive differences for both SGDR and ALES are below 10 cm,
which is in agreement with the high rate noise figures found in previous studies, such as
Zanife et al. (2003). In the Adriatic Sea, ALES is less noisy than the SGDR, an improve-
ment that becomes significant in J-2 196, due to the fact that most of the considered
short track passes over coastal waters, where SGDR performances are degraded, as seen
previously. Considering the tracks near South Africa, the situation changes because this
is an area where high significant wave height are often observed: the consequence is a
lower percentage of consecutive differences below 10 cm for both SGDR and ALES com-
pared to the tracks in the Adriatic Sea. ALES is slightly more affected by high SWH
conditions compared to the SGDR product in terms of noise, particularly for J-2 198.
This is also observable in the percentile interval, which is 3 cm wider for ALES in the
same track.

The percentile interval is significantly wider for SGDR in the Adriatic sea, which
confirms that the outliers detection performed on ALES successfully eliminates most of
the incorrect estimations. ALES mean and standard deviation of the distributions also
show values similar to SGDR, with an improvement for J-1 161 and J-2 196, which are
the two tracks where SGDR output showed the most significant degradation.

Biases between SGDR and ALES were also estimated for each track. Values were
averaged only at those along-track points where raw data showed a correlation with the
TGs higher than 0.9 for both SGDR and ALES TWLE and at least 50% of valid points
were available for both the datasets, therefore no bias was computed for J-2 196, where
no SGDR along-track points were so well correlated. Results are presented in table 5.
Biases are of the order of 1 cm, with standard deviations of the order of 2 to 5 cm. These
low values show how ALES constitutes a coastal-dedicated improvement of the standard
oceanic product without a significant loss of comparability with the SGDR estimation.

5. Conclusion

The present study aimed at the development and validation of ALES, the Adaptive Leading Edge Subwaveform retracker, which is capable of retrieving useful sea level information both in the open ocean and in the coastal zone. ALES estimates the parameters of interest while disregarding most of the trailing edge of a waveform, which is where deviations from the open ocean Brown model affect the quality of the standard altimetry product. The algorithm is based on a two-pass retracking which at first is only focused on the leading edge and then extends the estimation window including a variable number of gates depending on the SWH, in order try to ensure the same level of precision.

The retracker has been validated for Jason-1, Jason-2 and Envisat in two very different oceanographic areas. The results show on one hand that there is no significant degradation between standard open ocean high-rate TWLE estimation and ALES values in terms of noise and accuracy and on the other hand that ALES significantly increases the amount of altimetric information retrievable in the coastal zone, in areas that were disregarded even by dedicated coastal post-processed datasets, such as in the Gulf of Trieste.

It is very important to stress that the benefits are not limited to a 10 km wide strip along the coast, where the satellite footprint assimilates land return together with ocean signal, but extend to tens of kilometres from the coast. Deviations from the standard open ocean model are seen much further in the open sea, because land interference is not the main source of perturbation in the trailing edge. Bright targets are mainly connected to sea state and areas of low sea roughness that can be caused by land sheltering from winds or coastal upwelling fronts that produce film slicks on the surface.

ALES improves the quantity and the quality of coastal altimetry data without adding significant further noise to the estimation. Biases of the order of 1 cm with SGDR values demonstrate the comparability between the two datasets. Further analysis on bias and noise is needed and will be performed once an extensive global retracked dataset is available.

A dedicated quality control, independent from the comparison with SGDR, is needed in order to generate a consistent final product. Correlation and RMS analysis show that,
with proper removal of outliers, high-rate data are as accurate as 1-Hz post-processed sea
level heights. A quality criterion based also on the differences of consecutive estimates
can help the selection of data with low noise in order to avoid the need to average to
1-Hz, which limits the availability of measurements to one point roughly every 7 km.
Applications such as data assimilation into models and coastal circulation studies would
highly benefit from fine resolution coastal altimetry measurements.

The possibility to have a multi-mission retracker opens a wide range of further studies
in terms of applications and extensions. Validation with TG data also demonstrates that
ALES can be used for coastal sea level and surge monitoring studies, due to its increased
comparability with the coastal in-situ sensors. Recent studies such as Masina & Lamberti
(2013) demonstrate that there is interest in extreme sea level monitoring in the Northern
Adriatic, where ALES has been validated and could be used to improve the available
observations. ALES also retrieves SWH and σ⁰ and further research is needed to validate
these parameters against in-situ data.

The effort of the research on retracking should of course be assisted by new coastal-
dedicated corrections, in particular wet troposphere correction and sea state bias, and,
where possible, regional high resolution tidal models, in order to isolate the sea level
anomaly from the TWLE. Once SWH and σ⁰ are validated, sea state bias correction can
be improved using ALES estimates.

ALES should also be applied to previous (TOPEX,ERS) and present pulse-limited
altimetry missions (Alti-Ka, HY-2) to extend the time-series and realise the improvement
brought by the new altimeters and their smaller footprint.

Acknowledgements

We want to thank ISPRA (Istituto Superiore per la Protezione e la Ricerca Ambien-
tale) for providing sea level measurements from the tide gauge in Trieste. Special thanks
go to: Francesco De Biasio (CNR-ISMAR) for his help and assistance; Yang Le and Bao Li
Feng for sharing their algorithms; Luciana Fenoglio, Abderrahim Halimi, Walter Smith,
Luke West and Phil Woodworth for their suggestions. This work has been partially sup-
ported by the ESA/DUE eSurge (ESA/ESRIN Contract Number 4000103880/11/I-LG)
and eSurge-Venice (ESA/ESRIN Contract No. 4000104485/11/I-LG) projects.
References


Schumann, E. H. (1992). Interannual wind variability on the south and east coasts of


List of Figures

1 Characteristics of an idealised open ocean waveform. 35

2 Radargrams of Envisat track 416 over the Adriatic Sea (cycles 65 to 70). Land is shaded in grey. Every vertical line corresponds to a waveform at a latitude specified on the X axis. Y axis: gate numbers for each waveforms. The colorbar shows the power level for each gate. 36

3 The areas of study and the extent of the retracked altimetry passes. Top: Northern Adriatic and Gulf of Trieste. Bottom: Mossel Bay, along the South African coast. Bathymetry is shown by means of contour lines: they are drawn every 10 m. 37

4 Flow diagram of ALES retracking procedure for each waveform. Startgate and stopgate refer to the starting gate number and end gate number of the sub-waveform window. 38

5 Monte Carlo simulation (500 waveforms x gate x SWH value) using Envisat parameters. From upper left to lower right: RMSE difference between Brown full-waveform and subwaveform retracking at different SWH for A) Epoch, B) SWH and C) amplitude estimation; D) relationship between SWH and number of gates needed to estimate the Epoch with a 1 cm tolerance w.r.t. the full sub-waveform retracker. In A) the 1 cm tolerance is specified by a black horizontal line. 39

6 Examples of ALES retracking of Jason 2 waveforms for open ocean with SWH=0.75 m (left), coastal ocean with corrupted trailing edge and SWH=1.65 m (centre), open ocean with SWH=9.48 m (right). 40

7 North Adriatic: percentage of cycles highly correlated with TG time series (correlation coefficient ≥ 0.9) for Env 416 (top), J-1 161 (centre) and J-2 196 (bottom). On the x axis the along-track latitude of the nominal tracks are shown. Land is shaded in gray. The distance up to 20 km from the closest coastline is specified by a green line which refers to the y-axis on the right. For J-2 196, the latitude of the islet which separates the gulf from the lagoon is represented by a vertical dashed line. 41

8 South Africa: percentage of cycles highly correlated with TG time series (correlation coefficient ≥ 0.9) for Env 687 (top), and J-2 198 (bottom). On the x axis the along-track latitude of the nominal tracks are shown. Land is shaded in gray. The distance up to 20 km from the closest coastline is specified by a green line which refers to the y-axis on the right. 42

9 North Adriatic: percentage of outliers for Env 416 (top), J-1 161 (centre) and J-2 196 (bottom). On the x axis the along-track latitude of the nominal tracks is shown. Land is shaded in gray. The distance up to 20 km from the closest coastline is specified by a green line which refers to the y-axis on the right. 43

10 South Africa: percentage of outliers for Env 687 (top), and J-2 198 (bottom). On the x axis the along-track latitude of the nominal tracks is shown. Land is shaded in gray. The distance up to 20 km from the closest coastline is specified by a green line which refers to the y-axis on the right. 44
11 North Adriatic: correlation coefficient between Trieste TG and Env 416 (top), J-1 161 (centre) and J-2 196 (bottom). On the x axis the along-track latitude of the nominal tracks is shown and the latitude of the TG is highlighted with a black dot. Land is shaded in gray. The distance up to 20 km from the closest coastline is specified by a green line which refers to the y-axis on the right. For J-2 196, the latitude of the islet which separates the gulf from the lagoon is represented by a vertical dashed line. Negative values and correlation computed for less than 10 points in the time series are not shown.

12 South Africa: correlation coefficient between Mossel Bay TG and Env 687 (top), and J-2 198 (bottom). On the x axis the along-track latitude of the nominal tracks is shown and the latitude of the TG is highlighted with a black dot. Land is shaded in gray. The distance up to 20 km from the closest coastline is specified by a green line which refers to the y-axis on the right. Negative values and correlation computed for less than 10 points in the time series are not shown.

13 North Adriatic: absolute RMS difference between Trieste TG and Env 416 (top), J-1 161 (centre) and J-2 196 (bottom). On the x axis the along-track latitude of the nominal tracks is shown. Black dots on the x axis highlight the latitude of the TG (big dot) and of the closest point along track (small dot). Land is shaded in gray. The distance up to 20 km from the closest coastline is specified by a green line which refers to the y-axis on the right. A black dashed line marks the difference between the along track mean sea surface (DTU10) and the mean sea surface DTU10 at the TG. For J-2 196, the latitude of the islet which separates the gulf from the lagoon is represented by a vertical dashed line and.

14 South Africa: RMS of the difference between Mossel Bay TG values and Env 687 TWLE (top), and J-2 198 TWLE (bottom). The mean of the difference for each along track location was removed. On the x axis the along-track latitude of the nominal tracks is shown. Black dots on the x axis highlight the latitude of the TG (big dot) and of the closest point along track (small dot). Land is shaded in gray. The distance up to 20 km from the closest coastline is specified by a green line which refers to the y-axis on the right.

15 Histograms of consecutive TWLE differences in modulus for each track for SGDR (blue) and ALES (red). Columns of SGDR are thinner in order to facilitate the distinction.
Figure 1: Characteristics of an idealised open ocean waveform.
Figure 2: Radargrams of Envisat track 416 over the Adriatic Sea (cycles 65 to 70). Land is shaded in grey. Every vertical line corresponds to a waveform at a latitude specified on the X axis. Y axis: gate numbers for each waveforms. The colorbar shows the power level for each gate.
Figure 3: The areas of study and the extent of the retracked altimetry passes. Top: Northern Adriatic and Gulf of Trieste. Bottom: Mossel Bay, along the South African coast. Bathymetry is shown by means of contour lines: they are drawn every 10 m.
Figure 4: Flow diagram of ALES retracking procedure for each waveform. Startgate and stopgate refer to the starting gate number and end gate number of the sub-waveform window.
Figure 5: Montecarlo simulation (500 waveforms x gate x SWH value) using Envisat parameters. From upper left to lower right: RMSE difference between Brown full-waveform and subwaveform retracking at different SWH for A) Epoch, B) SWH and C) amplitude estimation; D) relationship between SWH and number of gates needed to estimate the Epoch with a 1 cm tolerance w.r.t. the full sub-waveform retracker. In A) the 1 cm tolerance is specified by a black horizontal line.
Figure 6: Examples of ALES retracking of Jason 2 waveforms for open ocean with SWH=0.75 m (left), coastal ocean with corrupted trailing edge and SWH=1.65 m (centre), open ocean with SWH=9.48 m (right).
Figure 7: North Adriatic: percentage of cycles highly correlated with TG time series (correlation coefficient $\geq 0.9$) for Env 416 (top), J-1 161 (centre) and J-2 196 (bottom). On the x axis the along-track latitude of the nominal tracks are shown. Land is shaded in gray. The distance up to 20 km from the closest coastline is specified by a green line which refers to the y-axis on the right. For J-2 196, the latitude of the islet which separates the gulf from the lagoon is represented by a vertical dashed line.
Figure 8: South Africa: percentage of cycles highly correlated with TG time series (correlation coefficient $\geq 0.9$) for Env 687 (top), and J-2 198 (bottom). On the x axis the along-track latitude of the nominal tracks are shown. Land is shaded in gray. The distance up to 20 km from the closest coastline is specified by a green line which refers to the y-axis on the right.
Figure 9: North Adriatic: percentage of outliers for Env 416 (top), J-1 161 (centre) and J-2 196 (bottom). On the x axis the along-track latitude of the nominal tracks is shown. Land is shaded in gray. The distance up to 20 km from the closest coastline is specified by a green line which refers to the y-axis on the right. For J-2 196, the latitude of the islet which separates the gulf from the lagoon is represented by a vertical dashed line.
Figure 10: South Africa: percentage of outliers for Env 687 (top), and J-2 198 (bottom). On the x axis the along-track latitude of the nominal tracks is shown. Land is shaded in gray. The distance up to 20 km from the closest coastline is specified by a green line which refers to the y-axis on the right.
Figure 11: North Adriatic: correlation coefficient between Trieste TG and Env 416 (top), J-1 161 (centre) and J-2 196 (bottom). On the x axis the along-track latitude of the nominal tracks is shown and the latitude of the TG is highlighted with a black dot. Land is shaded in gray. The distance up to 20 km from the closest coastline is specified by a green line which refers to the y-axis on the right. For J-2 196, the latitude of the islet which separates the gulf from the lagoon is represented by a vertical dashed line. Negative values and correlation computed for less than 10 points in the time series are not shown.
Figure 12: South Africa: correlation coefficient between Mossel Bay TG and Env 687 (top), and J-2 198 (bottom). On the x axis the along-track latitude of the nominal tracks is shown and the latitude of the TG is highlighted with a black dot. Land is shaded in gray. The distance up to 20 km from the closest coastline is specified by a green line which refers to the y-axis on the right. Negative values and correlation computed for less than 10 points in the time series are not shown.
Figure 13: North Adriatic: absolute RMS difference between Trieste TG and Env 416 (top), J-1 161 (centre) and J-2 196 (bottom). On the x axis the along-track latitude of the nominal tracks is shown. Black dots on the x axis highlight the latitude of the TG (big dot) and of the closest point along track (small dot). Land is shaded in gray. The distance up to 20 km from the closest coastline is specified by a green line which refers to the y-axis on the right. A black dashed line marks the difference between the along track mean sea surface (DTU10) and the mean sea surface DTU10 at the TG. For J-2 196, the latitude of the islet which separates the gulf from the lagoon is represented by a vertical dashed line and
Figure 14: South Africa: RMS of the difference between Mossel Bay TG values and Env 687 TWLE (top), and J-2 198 TWLE (bottom). The mean of the difference for each along track location was removed. On the x axis the along-track latitude of the nominal tracks is shown. Black dots on the x axis highlight the latitude of the TG (big dot) and of the closest point along track (small dot). Land is shaded in gray. The distance up to 20 km from the closest coastline is specified by a green line which refers to the y-axis on the right.
Figure 15: Histograms of consecutive TWLE differences in modulus for each track for SGDR (blue) and ALES (red). Columns of SGDR are thinner in order to facilitate the distinction.
### List of Tables

1. Data availability for this study. Row 2: the tide gauge to which the tracks specified in each column are compared. Row 3: Minimum distance between the tide gauge and the closest point on each satellite track. Row 4: number of reprocessed SGDR cycles; all the cycles where data in the area were available have been taken into account; the number refers to the number of cycles during which tide gauge data were also available. Row 5: number of CTOH cycles available. Row 6: number and percentage of 18 Hz (20 Hz) along track points where no waveforms from the SGDR of Envisat (Jason) were available. For Envisat, this also includes cases where waveforms were present, but the bandwidth in use was not 320 MHz. Row 7: number and percentage of along track points where no estimation of altimetric parameters was available in the SGDR, despite the availability of corresponding waveforms.

2. List of cycles. Column 2: list of the cycles used in this study for each satellite track. Cycles where satellite data were missing and/or coincident tide gauge data were not available have been excluded. The square brackets indicate the range of cycles including the extremes. Column 3: time interval in day/month/year format.

3. Validation results. Row 2: Outliers in absolute number and percentage of the total number of available waveforms. Row 3: median of the correlation of TWLE for each track with the correspondent tide gauge heights. Row 4 (rel): median value of the along-track relative RMS of the difference between TWLE and TG sea level height anomaly, with mean of the difference removed. Row 4 (abs): RMS of the difference of the absolute sea level heights above the ellipsoid at the closest point between TG and satellite tracks.

4. Criteria for valid points: minimum and maximum values of TWLE, SWH and $\sigma_0$.

5. Mean, standard deviation and percentiles (1st and 99th) of TWLE distributions for ALES (outliers removed) compared with SGDR. The total number of valid sea level observations for each track is reported in the last row.

6. Mean biases with standard deviation between SGDR and ALES TWLE estimations. Biases were computed using only along-track points where correlation coefficient with TG was higher than 0.9 and more than 50% of valid points were available for both SGDR and ALES. No such points were found for J-2 196, whose bias is therefore not computed.
Table 1: Data availability for this study. Row 2: the tide gauge to which the tracks specified in each column are compared. Row 3: Minimum distance between the tide gauge and the closest point on each satellite track. Row 4: number of reprocessed SGDR cycles; all the cycles where data in the area were available have been taken into account; the number refers to the number of cycles during which tide gauge data were also available. Row 5: number of CTOH cycles available. Row 6: number and percentage of 18 Hz (20 Hz) along track points where no waveforms from the SGDR of Envisat (Jason) were available. For Envisat, this also includes cases where waveforms were present, but the bandwidth in use was not 320 MHz. Row 7: number and percentage of along track points where no estimation of altimetric parameters was available in the SGDR, despite the availability of corresponding waveforms.

<table>
<thead>
<tr>
<th></th>
<th>Env 416</th>
<th>Env 687</th>
<th>J-1 161</th>
<th>J-2 196</th>
<th>J-2 198</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tide Gauge</td>
<td>Trieste</td>
<td>Mossel Bay</td>
<td>Trieste</td>
<td>Trieste</td>
<td>Mossel Bay</td>
</tr>
<tr>
<td>Minimum Distance</td>
<td>11.78 km</td>
<td>10.72 km</td>
<td>8.23 km</td>
<td>30.60 km</td>
<td>13.37 km</td>
</tr>
<tr>
<td>Number of SGDR cycles</td>
<td>57</td>
<td>58</td>
<td>99</td>
<td>165</td>
<td>126</td>
</tr>
<tr>
<td>Number of CTOH cycles</td>
<td>56</td>
<td>N/A</td>
<td>77</td>
<td>152</td>
<td>126</td>
</tr>
<tr>
<td>Missing waveforms</td>
<td>325 (2.71%)</td>
<td>0</td>
<td>1483 (4.77%)</td>
<td>5253 (15.45%)</td>
<td>1267 (2.53%)</td>
</tr>
<tr>
<td>Missing SGDR outputs</td>
<td>518 (4.32%)</td>
<td>384 (2.74%)</td>
<td>5966 (19.19%)</td>
<td>17653 (51.94%)</td>
<td>2981 (5.95%)</td>
</tr>
</tbody>
</table>
Table 2: List of cycles. Column 2: list of the cycles used in this study for each satellite track. Cycles where satellite data were missing and/or coincident tide gauge data were not available have been excluded. The square brackets indicate the range of cycles including the extremes. Column 3: time interval in day/month/year format.

<table>
<thead>
<tr>
<th></th>
<th>Valid cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Env 416</td>
<td>10 [12,14] 16 19 27 30 [42,45] 47 [49,52] [54,58] [60,72] [74,93]</td>
</tr>
<tr>
<td>Env 687</td>
<td>[14,17] [19,21] [23,26] 29 31 32 44 45 [47,50] [52,58] [62,73] [75,93]</td>
</tr>
<tr>
<td>J-1 161</td>
<td>[262,283] [285,292] [295,303] [306,340] [342,349] [351,363] [366,370]</td>
</tr>
<tr>
<td>J-2 196</td>
<td>[1,165]</td>
</tr>
<tr>
<td>J-2 198</td>
<td>[1,8] [12,14] [18,30] [32,35] [37,42] 45 46 [48, 80] 82 83 [85,91] [95,100]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Time interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Env 416</td>
<td>15/10/02 - 26/09/10</td>
</tr>
<tr>
<td>Env 687</td>
<td>13/03/03 - 7/10/10</td>
</tr>
<tr>
<td>J-1 161</td>
<td>16/02/09 - 23/01/12</td>
</tr>
<tr>
<td>J-2 196</td>
<td>19/07/08 - 31/12/12</td>
</tr>
<tr>
<td>J-2 198</td>
<td>19/07/08 - 26/06/12</td>
</tr>
</tbody>
</table>
Table 3: Validation results. Row 2: Outliers in absolute number and percentage of the total number of available waveforms. Row 3: median of the correlation of TWLE for each track with the correspondent tide gauge heights. Row 4 (rel): median value of the along-track relative RMS of the difference between TWLE and TG sea level height anomaly, with mean of the difference removed. Row 4 (abs): RMS of the difference of the absolute sea level heights above the ellipsoid at the closest point between TG and satellite tracks.

<table>
<thead>
<tr>
<th></th>
<th>Env 416</th>
<th>Env 687</th>
<th>J-1 161</th>
<th>J-2 196</th>
<th>J-2 198</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outliers</td>
<td>SGDR</td>
<td>ALES</td>
<td>SGDR</td>
<td>ALES</td>
<td>SGDR</td>
</tr>
<tr>
<td></td>
<td>1131 (9.43%)</td>
<td>1421 (11.85%)</td>
<td>384 (2.75%)</td>
<td>384 (2.74%)</td>
<td>8930 (28.72%)</td>
</tr>
<tr>
<td></td>
<td>2763 (8.89%)</td>
<td>20019 (58.90%)</td>
<td>3393 (6.78%)</td>
<td>3052 (6.10%)</td>
<td></td>
</tr>
<tr>
<td>Correlation</td>
<td>SGDR</td>
<td>ALES</td>
<td>SGDR</td>
<td>ALES</td>
<td>SGDR</td>
</tr>
<tr>
<td></td>
<td>0.44</td>
<td>0.88</td>
<td>0.95</td>
<td>0.95</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>0.93</td>
<td>0.87</td>
<td>0.44</td>
<td>0.44</td>
<td>0.94</td>
</tr>
<tr>
<td>RMS</td>
<td>SGDR</td>
<td>ALES</td>
<td>SGDR</td>
<td>ALES</td>
<td>SGDR</td>
</tr>
<tr>
<td></td>
<td>1.31 m (abs)</td>
<td>0.29 m (abs)</td>
<td>0.14 m (rel)</td>
<td>0.14 m (rel)</td>
<td>0.30 m (abs)</td>
</tr>
<tr>
<td></td>
<td>0.14 m (rel)</td>
<td>0.21 m (abs)</td>
<td>0.53 m (abs)</td>
<td>0.53 m (abs)</td>
<td>0.16 m (rel)</td>
</tr>
<tr>
<td></td>
<td>0.85 m (abs)</td>
<td>0.16 m (rel)</td>
<td>0.16 m (rel)</td>
<td>0.16 m (rel)</td>
<td></td>
</tr>
</tbody>
</table>
Table 4: Criteria for valid points: minimum and maximum values of TWLE, SWH and $\sigma^0$.

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>TWLE (m)</td>
<td>$-2 - \text{min TG value}$</td>
<td>$+2 + \text{max TG value}$</td>
</tr>
<tr>
<td>SWH (m)</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>$\sigma^0$ (dB)</td>
<td>7</td>
<td>30</td>
</tr>
</tbody>
</table>
Table 5: Mean, standard deviation and percentiles (1st and 99th) of TWLE distributions for ALES (outliers removed) compared with SGDR. The total number of valid sea level observations for each track is reported in the last row.

<table>
<thead>
<tr>
<th></th>
<th>Env 416</th>
<th>Env 687</th>
<th>J-1 161</th>
<th>J-2 196</th>
<th>J-2 198</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SGDR</td>
<td>0.3 cm</td>
<td>0.1 cm</td>
<td>-0.4 cm</td>
<td>-2.2 cm</td>
<td>0.0 cm</td>
</tr>
<tr>
<td>ALES</td>
<td>-0.4 cm</td>
<td>0.0 cm</td>
<td>-0.1 cm</td>
<td>-0.2 cm</td>
<td>0.1 cm</td>
</tr>
<tr>
<td>STD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SGDR</td>
<td>38.5 cm</td>
<td>10.1 cm</td>
<td>31.9 cm</td>
<td>47.6 cm</td>
<td>9.7 cm</td>
</tr>
<tr>
<td>ALES</td>
<td>31.4 cm</td>
<td>10.1 cm</td>
<td>15.6 cm</td>
<td>17.5 cm</td>
<td>10.5 cm</td>
</tr>
<tr>
<td>1% - 99%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SGDR</td>
<td>[-108.0, +94.7] cm</td>
<td>[-25.3, +25.5] cm</td>
<td>[-78.9, +70.5] cm</td>
<td>[-1.8, +99.8] cm</td>
<td>[-22.4, +22.1] cm</td>
</tr>
<tr>
<td>ALES</td>
<td>[-93.5, +77.9] cm</td>
<td>[-26.1, +26.1] cm</td>
<td>[-28.8, +29.4] cm</td>
<td>[-34.5, +28.5] cm</td>
<td>[-25.00, +25.2] cm</td>
</tr>
<tr>
<td>Obs Num</td>
<td>10569</td>
<td>13594</td>
<td>24394</td>
<td>15721</td>
<td>46662</td>
</tr>
</tbody>
</table>
Table 6: Mean biases with standard deviation between SGDR and ALES TWLE estimations. Biases were computed using only along-track points where correlation coefficient with TG was higher than 0.9 and more than 50% of valid points were available for both SGDR and ALES. No such points were found for J-2 196, whose bias is therefore not computed.

<table>
<thead>
<tr>
<th></th>
<th>Env 416</th>
<th>Env 687</th>
<th>J-1 161</th>
<th>J-2 198</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias TWLE (cm)</td>
<td>-1.3 ± 3.5</td>
<td>-1.7 ± 2.6</td>
<td>-1.2 ± 4.4</td>
<td>-0.3 ± 3.4</td>
</tr>
</tbody>
</table>