Project information

Project information	
Project full title	EuroSea: Improving and Integrating European Ocean Observing and Forecasting Systems for Sustainable use of the Oceans
Project acronym	EuroSea
Grant agreement number	862626
Project start date and duration	1 November 2019, 50 months
Project website	https://www.eurosea.eu

Deliverable information

Deliverable information	
Deliverable number	D5.7
Deliverable title	**Automated tide gauge data quality control software and report**
Description	Software that will be provided to all stakeholders to perform automated QC of tide gauge data
Work Package number	5
Work Package title	Coastal Resilience and Operational Services Demonstrator
Lead beneficiary	National Oceanography Centre (NOC),
Lead authors	Angela Hibbert, Jue Lin, Begoña Pérez Gómez
Contributors	
Due date	31/08/23
Submission date	31/08/23
Comments	This deliverable contributes to Task T5.1.1 Low cost and maintenance free tide gauges, of which the NOC is the primary beneficiary. However, delivery of the QC work has been subcontracted by Puertos del Estado (PdE) to Nologin by grant amendment.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 862626.

https://doi.org/10.3289/eurosea_d5.7

Table of contents

Executive summary

The WP5 Coastal Resilience and Operational Services demonstrator aims to design and deploy innovative sea level monitoring systems, integrating their observations with downscaled model forecasts into an alert-based monitoring and forecasting tool (OSPAC) that is design for use by ports and local authorities. As part of this data integration process, sea level observations must be quality-controlled in near real-time, to minimise the risk of false alarms. This has historically been achieved using an established open-source software package (SELENE). Recent progress using automatic quality control (QC) in delayed mode has led to the development of additional functionality that could enhance the SELENE software. At the same time, the delivery of new sea level time series from Global Navigation Satellite Systems (GNSS), provides an additional means of data validation. WP5 aims to enhance the SELENE software by incorporating these enhancements and new data feeds, thereby improving the quality of the OSPAC tool. A delay in the installation of the sea level monitoring systems has led to a corresponding delay in the provision of data inputs to SELENE and subsequently to OSPAC. Some development work has been possible using alternative test time series, but a key implementer in presently taking parental leave, which will delay further development work until October 2023. This deliverable report will be updated thereafter.

1. Introduction

WP5 aims to demonstrate an end-to-end solution for sea level observing and forecasting, in what has been termed the Coastal Resilience and Operational Services demonstrator. Commencing with the design and deployment of innovative prototype sea level monitoring systems by NOC, the collected observations are then to be assimilated (alongside downscaled ocean analyses and model forecasts) into a bespoke operational monitoring software tool for coastal ports and cities produced by WPS delivery partners Puertos del Estado (PdE) and Nologin. The software tool, known as OSPAC (Oceanographic Services at the Service of Ports and Cities), is to provide an alert-based monitoring and forecasting system for port authorities and local government bodies in 3 target locations : (1) Barcelona, Spain, (2) Taranto, Italy and (3) Buenaventura, Colombia.

Since OSPAC is to provide an alert-based monitoring and forecasting system for each location, it is essential to embed near real-time data quality control functionality in order to prevent the occurrence of false alarms. This has initially been achieved by incorporating well-established QC software devised by PdE known as SELENE (SEa LEvel NEar-real time quality control processing) into the OSPAC software. However, Task 5.2 (Oceanographic Services at the Service of Ports and Cities) also aims to update the SELENE software to provide extra functionality and thereby improve the quality of the sea level observations. These proposed additional utilities include refined interpolation and spike detection algorithms and comparisons with data from a new sea level monitoring technique known as Interferometric Reflectometry (IR). This technique can generate sea level observations from the 3 new GNSS receivers that are being installed as part of the tide gauge instrumentation at the 3 locations and these additional sea level observations can then be used for datum control of the primary sea level measurement technology. In other words, these additional data streams will improve the early detection of datum shifts, spikes and other important data features, thus allowing users to identify system faults and reduce the risk of issuing false alerts.

WP5 also aims to train each set of local stakeholders in the use of the automated quality control software as well as in the use of OSPAC itself (as described in deliverable D5.6 Documentation associated to the capacity building).

The implementation of the updated SELENE quality control software is contingent upon the timely provision of near real-time data streams from the new sea level monitoring systems that were planned under Task 5.1.1. Unfortunately, these tide gauge installations were delayed considerably by multiple factors including the global COVID-19 pandemic, issues relating to the UK departure from the European Union and the conflict in Ukraine. Consequently, new data streams were, until very recently, unavailable for integration within the OSPAC software, delaying the enhancement to the SELENE software.

However, so as not to lose momentum entirely, the new functionality was tested using the GNSS-IR technique on existing locations for which it was known to work (the REDMAR tide gauges at Barcelona and Tarragona). This allowed some development of the SELENE routines to take place in OSPAC, prior to completion of the new EuroSea tide gauge installations in Barcelona, Taranto and Buenaventura. Nevertheless, the considerable delay in providing the new data streams has meant that that new functionality has not yet been implemented in a live environment, but is planned for development in October 2023. This deliverable report therefore describes the use of SELENE software at present, the proposed enhancement with GNSS-IR and other functionality and an example of the added-value that this provides using another Mediterranean tide gauge. This deliverable will be updated once the enhancement to the software is complete.

2. Enhancement of SELENE software

2.1. Current functionality of SELENE software

The SELENE software was originally developed at PdE in the 1990’s as a FORTRAN-based software tool that was integrated into their NIVMAR sea level forecasting system, but was recently upgraded by PdE to a well-documented open source python-based software tool that is accessible via github (https://puertos-del-estado-medio-fisico.github.io/SELENE/). It is run operationally on the Spanish REDMAR network of sea level stations and is intended to perform near real-time automatic data QC procedures to minimise the need for human intervention. These QC procedures meet the requirements for L1 data quality that are demanded by organisations such as the Global Sea Level Observing System (GLOSS).

The SELENE software consists of a number of key modules which are invoked and applied in accordance with the data flow shown in Figure 1. These modules comprise:

QC module

The QC module takes as its input the raw sea level time series at the original sampling intervals and applies tests for out-of-range values, the detection of spikes, stuck data (i.e. data that remain at a constant value), as well as checking for timing errors and that data are in the expected format. Data flags are then applied and the ‘flagged’ data are output at the original sampling interval.

Interpolation module

The interpolation module interpolates small gaps (of less than 10-25 mins) in the flagged data file and also interpolates small sections of data that have been flagged as suspect. This module also resamples the original data to 5 min intervals.

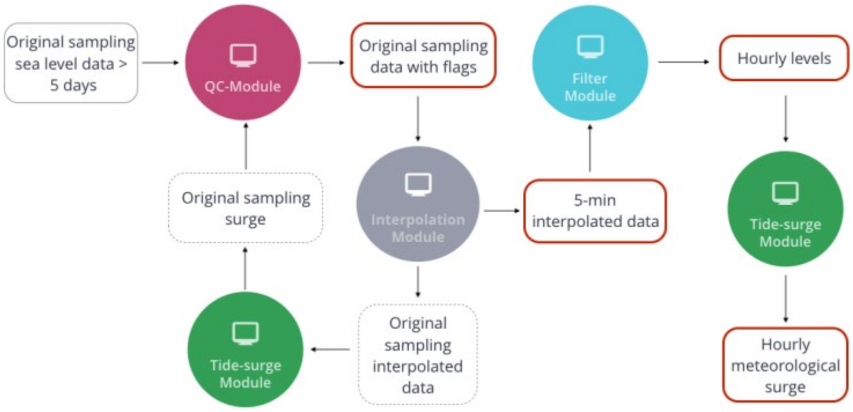

Figure 1: SELENE modules (QC, Interpolation, Tide-Surge and Filter) and data flow of the software. The final output products are framed in red and consists of the original sampling data with added data flags, 5-min interpolated data, hourly sea levels and residuals (aka meteorological surge component) (Pérez-Gómez et al, 2013).

Tide-surge module

This module uses tidal constituents (derived from the harmonic method of tidal analysis) to reconstruct the tidal sea level components of the time series at the original sampling interval and to separate it from the non-tidal (residual) sea level component. Since the tidal component often dominates a sea level record, this allows underlying errors to be identified more easily in the residual time series and the QC and interpolation modules can be applied iteratively, to produce quality-controlled data at 5 min and original sampling intervals. The tide-surge module is also invoked to separate the tidal and non-tidal (surge) components once data have been filtered to hourly intervals, outputting files of hourly non-tidal residuals.

Filter module

The filter module applies the Pugh (1987) filter to the quality-controlled 5 min data files, producing files of hourly total water level data.

Thus, the current version of SELENE outputs 3 quality-controlled datasets: (1) at the original sampling interval, (2) at 5 min sampling and (3) at hourly sampling

2.2. Proposed enhancements to the SELENE software

In 2019, the National Oceanography Centre (NOC) produced an alternative automatic QC software package (Williams et al., 2019) that was designed to be used for delayed-mode sea level data processing by organisations that lack the expertise to perform manual QC. This software package took its data input from the IOC's Sea Level Station Monitoring Facility (https://www.ioc-sealevelmonitoring.org/) which is a global data portal for near real-time sea level observations. The NOC QC software made use of some elements of the SELENE software, particularly the spike detection algorithm, but demanded a more flexible system that was suitable for a wider range of tide gauge data than was encountered in Spain. For example, the SELENE software requires an additional data input comprising tidal constituents derived from harmonic analysis of a 12-month sea level record for each location, whereas users of the NOC QC software would not have the necessary expertise to produce this. Consequently, the NOC QC software had in-built tidal analysis functionality, which enabled the fitting and removal of the tide at an earlier stage in the process, without prior knowledge of the tidal regime. Since this tidal fitting routine works on time series that contain gaps, this avoids the need to interpolate missing data at an early stage. If gaps must be filled for later processing, a better estimate of the missing data is instead obtained by interpolating the non-tidal residual instead of total water level.

Other refinements in the NOC QC software included:

- An improved spike detection algorithm that uses median rather than mean statistics and is more robust at identifying clusters of data spikes.
- A new algorithm for stuck oscillatory data (i.e. data that vary occasionally from an otherwise constant value, as shown in the example in Figure 2).
- Channel comparison functionality allowing: (1) automatic identification of discrepancies between data from different parallel instruments (referred to as data channels) and (2) selection of the best of 2 or 3 data channels, even where they are of different instrument types (for example, comparing 5 min sampling float gauge sensors with 1 min sampling radar sensors).

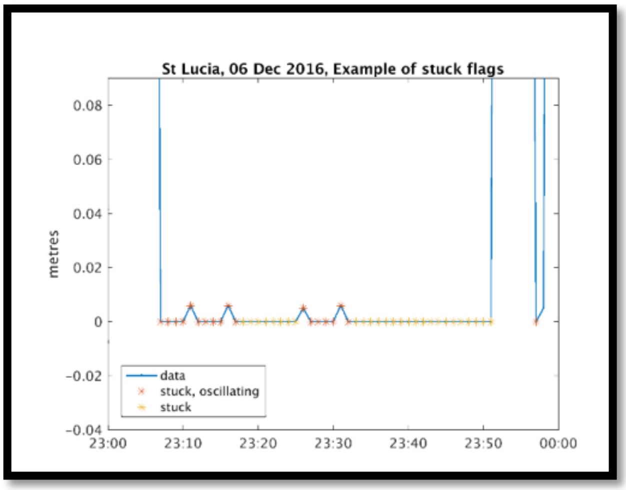

Figure 2. An example of stuck oscillatory sea level data

In contrast to SELENE, which is open-source software and is designed to be implemented in near real time, the NOC software is Matlab-based and built to be used in delayed mode. Whilst the use of the NOC QC software does not demand that the user should hold a Matlab software licence (as the QC software can be run as an executable command), conversion to open source software would undoubtedly render it more accessible to other users. Consequently, as part of Task 5.2, it is proposed that some of the enhancements contained in the NOC delayed-mode automatic QC software will be considered for incorporation into the EuroSea update to the SELENE near real-time automatic QC software. In particular, the new channel comparison functionality, will facilitate intercomparisons of the time series from the conventional sea level sensors and those that are generated by the GNSS-IR technique.

GNSS-IR

GNSS receivers are conventionally used to measure changes in land elevation by detection of positioning and timing information from a constellation of navigational satellites such as GPS (USA), and GALILEO (Europe). As such, they are a valuable means of detecting land motion trends that contribute to the long-term sea level trends observed by tide gauges, so the 2 systems are often co-located. The strength of the GNSS signal at a receiver (and therefore the quality of the positions) can be affected by the local surroundings, as a direct GNSS signal and its reflection off a nearby object will interfere. Where the signal is reflected off a relatively flat surface like the sea, it turns out that the combination of the direct and reflected signal (the signal to noise ratio) will exhibit a periodic variation with a frequency that varies according to the elevation of the satellite and the level of the sea surface, thus allowing sea level height to be inferred. Currently, this interferometric reflectometry (IR) technique does not afford the high frequency sampling (<1 min) or latency of communications (<6 min) that is ideal for tsunami monitoring, but using the EuroSea GNSS receivers at Barcelona and Taranto (shortly to be followed by Buenaventura) observations at intervals of 15 min and a latency of 1 hour have been produced. The WP5 team therefore proposes to use these new time series as an additional means of error detection within the channel comparison functionality of SELENE. An example of this application is given below.

Example of the channel comparison technique using the NOC Gibraltar GNSS

The NOC has maintained a tide gauge in Gibraltar for several decades. The current instrumentation comprises dual underwater pressure sensors and an above water radar sensor, with a co-located GNSS receiver for monitoring vertical land motion. In July 2022, a local contact in Gibraltar reported that a cone that guides the radar pulses from the radar sensor towards the sea surface appeared to have been broken off. As a result, it was expected that the data from the sensor would be unreliable until a repair could be made. However, when comparisons of the data from the radar sensor were made with those from the underwater pressure sensors (PR1 and PR2), it was apparent that there were possible errors in several of the time series. When the radar times series and that of PR1 were differenced (Figure 3), there appeared to be both a systematic drift between them since July 2022 as well as an increase in noise since March 2020. PR2 did not display the systematic trend exhibited by PR1 since July 2022, but it was unclear whether this trend was real or instrumental. Using GNSS-IR, an additional sea level time series was derived from the Gibraltar GNSS and this was consistent with the record from PR2. This intercomparison (Figure 3) showed that PR1 was the source of the systematic drift since July 2022, whilst the radar sensor was the source of the noise that arose since March 2020 and it was likely that the radar cone damage had been sustained at that point. The time series from PR2 and the GNSS receiver were therefore deemed to be the most robust and were used in place of the other sensors pending repairs.

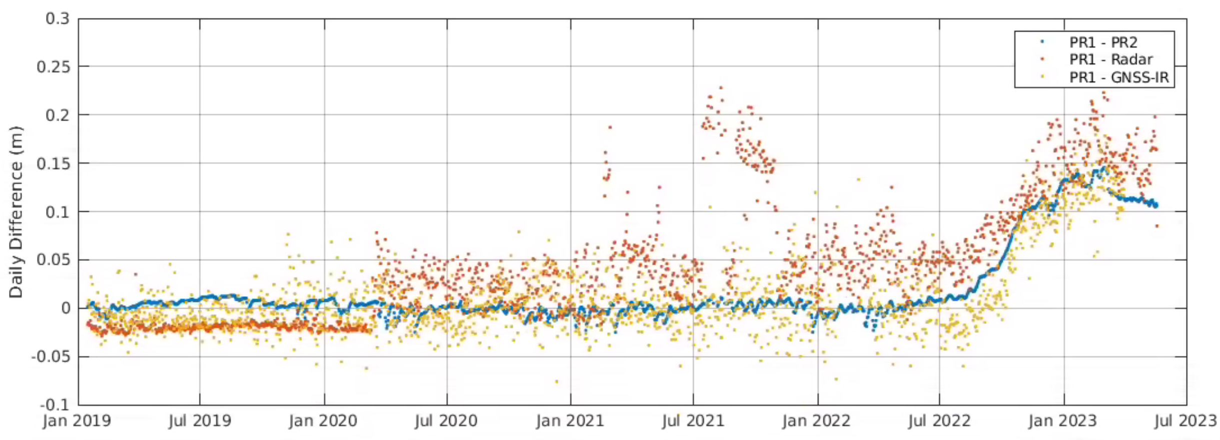

Figure 3. An example of data channel intercomparison (by differencing) for the Gibraltar tide gauge. PR1 denotes pressure sensor 1, PR2 denotes pressure sensor 2.

It is anticipated that the planned channel comparison functionality in the SELENE software will automate this validation process.

2.3. Progress to date

Since the installation of the EuroSea tide gauges was subject to significant delays, the new GNSS-IR sea level time series have only recently become available for incorporation within the SELENE software, which has limited the extent of work that could be undertaken pre-installation. However, to make some headway, the WP5 implementers have taken the step of testing the incorporation of a GNSS-IR channel comparison function using data from the Barcelona and Tarragona GNSS receivers that are collocated with the REDMAR tide gauges at these sites. This has helped to expedite development of software in preparation for the delivery of the live EuroSea GNSS-IR data feeds.

A key member of the WP5 delivery personnel who are responsible for delivery of the upgraded SELENE software is presently taking parental leave, so the remaining upgrade works will be delivered on their return which is likely to take place in October 2023. This deliverable will be updated accordingly once that work is complete.

3. Conclusion

The applications of automatic quality control software have recently been extended from near-real time observations to delayed-mode data, which has resulted in the development of improved functionality. Task 5.2 aims to incorporate some of these enhanced functions into the SELENE QC software as part of the Coastal Resilience and Operational Services demonstrator, whilst also including an additional source of datum control and error identification from a GNSS-IR time series of sea level. It is anticipated that this will further improve the resilience of the OSPAC forecasting tool and will minimise the risk of false alerts. This deliverable report will be updated once these enhancement are complete.

4. References

Pérez-Gómez, B., Alvarez Fanjul, E., Pérez, S., de Alfonso, M. and Vela, J., 2013. Use of tide gauge data in operational oceanography and sea level hazard warning systems. Journal of Operational Oceanography, 6(2), 1–18.

Pugh, D. (1987) Tides, Surges and Mean Sea Level: A Handbook for Engineers and Scientists. John Wiley & Sons, Chichester, 472 p.

Williams, J., Matthews, A. and Jevrejeva S, 2019, Development of an Automatic Tide Gauge Processing System, NOC Research & Consultancy Report No. 64, NOC_RandC_64_Final.pdf (psmsl.org)