https://doi.org/10.5194/egusphere-2023-733
Preprint. Discussion started: 20 April 2023

Delayed-mode reprocessing of in situ sea level data for the Copernicus Marine Service

Jue Lin-Ye[1,*], Begoña Pérez-Gómez[2], Alejandro Gallardo[1], Fernando Manzano[2], Marta de Alfonso[2], Elizabeth Bradshaw[3], Angela Hibbert[3]

[1] Nologin Ocean Weather Systems. Paseo de la Castellana 216, Planta 8ª, Oficina 811, 28046 Madrid, España

[2] Puertos del Estado. Avenida del Partenón 10, 28042, Madrid, España

[3] National Oceanography Centre. Joseph Proudman Building, 6 Brownlow Street, Liverpool L3 5DA, United Kingdom

[*] *Correspondence to*: Jue Lin-Ye (jue.lin@nologin.es)

Abstract. The number of tide gauges providing coastal sea level data has significantly increased in recent decades. They help in the issue of coastal hazard warnings, the forecasting of storm surges and tsunamis, and in operational oceanography applications. This data is often automatically quality controlled in near-real-time, as is the case in the Copernicus Marine Service. However, machine and visually inspected validation of historical sea level time series from these platforms is required. It has long been requested by the oceanographic community at large. As a result, a new initiative seeks to provide such validated data for the Copernicus Marine Service, by developing a new delayed-mode reprocessed in-situ sea level product. The existing software for the near-real-time product has been adapted to this end. After being visually controlled by a team with expertise in sea level reprocessing, the new product was launched in November of 2022. The entire reprocessing is discussed, in detail. An example of the information that can be extracted from the delayed-mode product is also given.

Keywords: sea level, observation, delayed-mode, software

1 Introduction

Sea level has long been in the public spotlight for being one major indicator of Climate Change and is now an essential ocean variable (https://www.goosocean.org/index.php?option=com_content&view=article&id=14&Itemid=114l). Remote and in-situ observations show that the sea level rise has been accelerating in recent decades (Nerem et al., 2018; Dangendorf et al., 2019). It raises concern about a corresponding increase in the magnitude, frequency and impact of coastal extreme events (Vousdoukas et al., 2018; Kirezci et al., 2020; Almar et al., 2021).

As one of the cheapest, most readily monitored ocean variables, it is of no surprise that it was one of the first ones to be systematically measured. Tide gauge data are an essential source of information to understand the contribution of different physical processes to sea level variations at different timescales. Therefore, reliable delayed-mode reprocessed sea level data are crucial for validation of hindcasting and forecasting ocean models, for both short-term forecasts and long-term climatic studies.

https://doi.org/10.5194/egusphere-2023-733
Preprint. Discussion started: 20 April 2023

The Copernicus Marine Service gathers the information of a network of tide gauges with other in-situ observations (Copernicus Marine In Situ Tac Data Management Team, 2021). Until November of 2022, these platforms were only included in the near-real-time data product, while a sea level delayed-mode reprocessed product was pressingly requested for model validation, comparison to remote-sensing measurements, among other functions.

There is, currently, a contract under way to deliver this delayed-mode reprocessed sea level product for the Copernicus Marine Service (Product User Manual, 2022; Quality Information Document, 2022). The spatial coverage consists mostly of European waters, at this moment (Fig. 1). The most veteran tide gauges are in the Baltic and the Iberian-Biscay-Irish regions. Some of these stations still operate and offer historical series of over 100 years. Most data in the Mediterranean Basin had series shorter than one year, as of 2020. New tide gauges are emerging from Turkey and other Eastern Mediterranean countries. Some geographically non-European islands also contribute at the global level to this product, which will expand to include other stations of the Global Sea Level Observing System (GLOSS) and improve the global coverage by November 2024.

Stations in different regions experience different sea level variability (Pugh, 1987). Astronomical tides differ from one location to the next, depending on the extension of the sea basin and the degree of openness of the waters. For instance, the Mediterranean Sea has a relatively reduced semi-enclosed basin, with a smaller astronomical tidal component. On the contrary, the Atlantic coastlines are exposed to the open ocean and have a strong astronomical tidal component. Each region also experiences different atmospheric conditions, which influence storm surges. There can be meteo-tsunamis, seiches and other high frequency events, as well. Thus, each station can register a variety of total sea levels, astronomical tides and de-tided sea levels.

The time series used in this product span very different periods, ranging from 1886 in some Baltic stations, to dozens of stations that were installed as late as 2019 (Fig. 2). The newer sensors have different precisions, which tend to be in millimeters, compared to older sensors, used at some point in the century-long series, that can have a factual precision of centimeters. Following the Baltic area, more tide gauges were successively installed in the Iberian-Biscay-Irish region, the Mediterranean region, the Arctic, the North-West shelf, and more recently, the Black Sea region (Fig. 2 and 3).

Up to now, the quality control of in-situ measurements has relied heavily on expert visual inspection. This is still a key part of the process, but the community has increasingly adopted automation to detect low quality data, more quickly and exhaustively. Likewise, expert knowledge, ruled by physical formulation, can be partially translated into machine language. The combination of automation and an efficient algorithm can minimize the total cost of the quality control, and this can make it possible to apply quality control to a large network of thousands of tide gauge stations (Eurosea project. Deliverable 3.3).

2 The near-real-time: The near-real-time (L1) product and the SELENE software

The near-real-time process (Fig. 4) in Copernicus takes total observed sea levels at the original time sampling, applies an automatic quality control test, and flags sea level data points, where appropriate. These series are transformed into an intermediate, 5-min sampled format, where short gaps (below or equal to 27 minutes) are interpolated. Then, a Pugh filter (Pugh, 1987) is applied to produce hourly sea levels. In the Copernicus Marine Service, the original, automatically flagged total sea level and the hourly filtered total sea level are both available as near-real-time products.

The quality control of tide gauge time series is performed with the SELENE software (Puertos del Estado, 2019). The original software was created in the 1990s, based on Fortran77 programming language and C-Shell scripts. It was implemented, operationally, in near-real-time (every 15 min). It included quality control, filtering to hourly values and computation of non-tidal residuals. These de-tided sea level observations facilitate quality control of the total sea level time series, since tidal variations often dominate a sea level record, making subtle

https://doi.org/10.5194/egusphere-2023-733
Preprint. Discussion started: 20 April 2023

errors more difficult to detect. The corrected time series were then used to improve the sea level forecasts of the Nivmar Sea Level Forecasting system of Puertos del Estado (Álvarez-Fanjul et al, 2001; Pérez-Gómez et al., 2013). This was performed by means of a nudging technique, still operational today.

The interest of international tide gauge communities (such as the Copernicus Marine Service, GLOSS and EuroGOOS Tide Gauge Task Team) in applying this automatic algorithm to other tide gauge networks resulted in the software being upgraded to a modern, open-source (Python) and well-documented format. It is multi-platform (Linux, MacOS and Windows) and it is now used by some regions in the Copernicus Marine Service In Situ TAC, for purposes outside the Copernicus system.

The SELENE software follows best practices, as defined by GLOSS (IOC, 2020). In near-real-time, it applies the L1 quality control level, as defined by the IOC (2020), including: out-of-range, spike detection, and stability tests. It also includes a module for de-tiding, based on the Foreman tidal prediction package in Fortran77 (Foreman, 1977). However, this module is not yet applied in operational near-real-time mode in the Copernicus Marine Service In Situ near-real-time product, as it requires historic tidal constants at each station, as a precursor to the process.

The SELENE software uses a series of station-configuration parameters. These are the maximum and minimum total sea level; degree of polynomial and other parameters for spike detection; parameters for the stuck-data test (data that keeps being constant for an abnormally long time), among others. Prior to this project, no study was carried out to calibrate the parameters for each station. Such a study needs to employ statistical analyses, while dealing with the physical oceanographic aspects.

3 The delayed-mode: the reprocessed product (L2) and the methodology

The delayed-mode reprocessing follows the same procedure as the near-real-time process. The products (https://doi.org/10.48670/moi-00307) available in the first release of November 2022 comprise quality-controlled versions of the original series of total sea level and the hourly filtered series of sea level. These end in December 2020, and are available for the 639 stations that were operational by that time. The series that by December of 2020 presented less than one year of data were excluded (Fig. 3).

An updated version of the SELENE software is used for quality control. The reprocessing is complemented by visual inspection. The updated SELENE software contains a series of new modules. The changes can improve near-real-time quality control as it includes new quality control tests and quality control of non-tidal residuals (by providing yearly tidal constants). A new version of the de-tiding module had been incorporated, in Python language. The corresponding Python package is called uTide (Codiga, 2011), also based upon Foreman. It is used by other authors, such as Mélet et al. (2021), for tide gauge data reprocessing. uTide accepts non-uniform sampling frequencies. This flexibility comes in especially handy, since some time series uploaded to Copernicus Marine Service are not entirely uniform, in terms of sampling interval. Apart from being a product itself, the de-tided sea level can help detect anomalous behaviour, such as in Fig. 5. These small errors are not easily detectable by visual inspection, and this automation can significantly reduce workload. Other issues detectable with de-tided sea levels are clock malfunction and datum changes, when combined with visual inspection. Hourly non-tidal sea level for the Iberian-Biscay-Irish region will be available by November 2023.

One of the new modules added to SELENE software is the "buddy" checking or neighbour test, applied to hourly de-tided series and monthly mean sea levels. It has been implemented to compare the target station to neighboring stations within a range of 0.01º, approximately 1 km. Another new module is dedicated to identification of data displaying an erroneous attenuation of the tidal signal. Specific testing had been carried out, to set the criterion for attenuated data, in contrast to neap tides (Product User Manual).

https://doi.org/10.5194/egusphere-2023-733
Preprint. Discussion started: 20 April 2023

Finally, an exhaustive study was carried out, to calibrate the parameters for SELENE. Specifically, a sensitivity test has been performed, and the goodness of the results has been controlled by expert visual inspection. This process has been repeated for each one of the 909 files of various sampling frequencies, from the 639 stations in the Copernicus Marine Service In Situ TAC. The selection of the parameters followed the application of unified criteria. For instance, the limit for stuck data depends on the precision of the data. These customized parameters can improve the performance of SELENE by successfully detecting spikes, even in difficult cases (Fig. 6).

The processing speed of the first Python version of the SELENE algorithm was slower than that of the original Fortran77 code. It hampered the reprocessing of extensive historical series. During this project, the computational cost of the Python-based algorithm has been reduced ten-fold. In addition, the reprocessed product has been assigned with corresponding metadata, to better identify the product from each station. Product enhancements are planned such as improvement of the geographical coverage, by including stations from the GLOSS network in 2024. Time extensions of the product to six months prior to the respective release will be regularly provided. Future goals for the reprocessing service include product validation with satellite altimetry, and to carry out a feasibility study on the inclusion of vertical land motion corrections, based on Global Navigation Satellite System (GNSS) receivers.

4 Possible applications

The reprocessed total sea level product has many applications for scientists and coastal stakeholders. As an example, seasonal extreme water levels can be computed, normalized (divided) by the tidal range. Seasonal relative maximum sea levels derived in this way could help highlight regions at most risk of flooding. Alternatively, seasonal relative lowest water levels can also be useful to remind harbours to keep their maintenance on schedule, to ensure its channels have enough draft for ships to circulate freely.

From hourly (filtered) total sea level, the maximum and minimum values for each season are computed. The values for all stations are scaled, to share a single zero. These extreme sea level parameters are subsequently divided by the tidal range at each station. The tidal range is computed as the difference between the historical 99th percentile and the historical 1st percentile. The relative maximum and minimum hourly total sea levels in each season are shown in Figs. 7 and 8, respectively. Note that the used series are the ones that by December of 2020 presented a time interval longer than one year. Moreover, each series has a different length, as shown in Fig. 3.

The (absolute) highest water levels, according to the European Environment Agency and to the reprocessed sea level data, are in the Iberian-Biscay-Irish area and the Northwest-shelf. The areas with the most extreme water levels are the British coasts in the Irish Sea, as well as the French coast in the English Channel. Some high-water levels are typical in the Frisian Sea, near the German-Denmark border. This area, too, is heavily influenced by the Atlantic Sea. However, these regions are geographically and technologically more prepared for such ranges of sea level, since they are typically macro-tidal. Attention must be paid to micro-tidal and meso-tidal areas with large percentage increases in sea level, for small changes can induce floods in them.

The Strait of Kattegat (between the east coast of Denmark and the South-western coast of Sweden, 56.82ºN, 11.39ºE) suffers from the highest relative sea levels in Spring, Autumn and Winter (Fig. 7). Furthermore, during the Autumn-Winter, the phenomenon of high relative sea levels extends to the neighboring Gulf of Bothnia (the gulf between Sweden and Finland, 62.33ºN, 19.55ºE). The Strait of Kattegat also presents the lowest relative sea levels in Autumn and Winter (see Fig. 8). The North-western Mediterranean displays high relative sea levels in Spring, Autumn and Winter. It is known that some of these regions, like Barcelona, in the North-western Mediterranean, have a history of engineering challenges for coastal protection, like during Storm Gloria (Pérez Gómez et al., 2021). Similarly, Denmark and Sweden had taken steps to improve their protection against floods, because of the historical flood in 1872 (Hallin et al., 2021; Fredriksson et al., 2018). The improvements were not always based on engineering, but also on changes in the structure of the economic system, insurance policies, among others. Much more information can be drawn from the reprocessed sea level data. All this will improve communication between stakeholders.

5 Conclusions

The Copernicus reprocessed sea level product is especially useful to the oceanographic community. The quality of the data can determine both the veracity of the conclusions and the correctness of the decisions to be taken. The delayed-mode reprocessing is based on the near-real-time process, the existing software and visual inspection for quality control. There were several improvements and adaptations in the automated part, such as the adaptation of the de-tiding module from Fortran77 to Python language, the implementation of the buddy checking, the detection of attenuated data, and the calibration of parameters for the SELENE software.

An example of a possible use for these reprocessed sea level data is given. It helps map the risk areas for high and low relative sea levels, per season of the year, in Europe. Temporal and geographical extensions are planned for the present service. There will also be improvements for the de-tiding module, tests on comparison with altimetry and inclusion of Global Navigation Satellite System land vertical motion corrections.

Competing interests

The contact author has declared that none of the authors has any competing interests.

Acknowledgements

We use the uTide python package in the upgraded SELENE software. Many thanks to the developers for their help solving questions raised during the implementation of the tool.

Bibliography

Almar, R., Ranasinghe, R., Bergsma, E.W.J., Diaz, H., Mélet, A., Papa, F., Vousdoukas, M., Athanasiou, P., Dada, O., Almeida, L.P. and Kestenare, E. (2021). A global analysis of extreme coastal water levels with implications for potential coastal overtoppping. Nature Communications, (2021), 12(1), doi:10.1038/s41467-021-24008-9.

Álvarez Fanjul, E., Pérez Gómez, B. P., & Arévalo, I. R. S. (2001). Nivmar: a storm surge forecasting system for Spanish waters. *Scientia Marina*, *65*(S1), 145-154.

Codiga, D. L. (2011). Unified tidal analysis and prediction using the UTide Matlab functions.

Copernicus Marine In Situ Tac Data Management Team (2021). Product User Manual for multiparameter Copernicus In Situ TAC (PUM). https://doi.org/10.13155/43494

Dangendorf, S., Hay, C., Calafat, F. M., Marcos, M., Piecuch, C. G., Berk, K., & Jensen, J. (2019). Persistent acceleration in global sea-level rise since the 1960s. *Nature Climate Change*, *9*(9), 705-710.

European Environment Agency. https://www.eea.europa.eu/data-and-maps/figures/mean-tidal-amplitude [Retrieved: 15-02-2023]

https://doi.org/10.5194/egusphere-2023-733
Preprint. Discussion started: 20 April 2023

Eurosea project. Deliverable 3.3. *New Tide Gauge Data Flow strategy.* https://eurosea.eu/download/eurosea-d3-3-new-tide-gauge-data-flow-strategy/?wpdmdl=3584&refresh=63eb77aad21ea1676375978

Foreman, M. G. G. (1977). Manual for tidal heights analysis and prediction. Pacific Marine Science Report 77-10. *Institute for Ocean Sciences, Sydney, Canada*.

Fredriksson, C., Martinez, G., Larson, M., & Eellend, B. F. (2018). Using historical storms for flood risk management: the 1872 storm in south Sweden. *RCC Perspectives*, (3), 11-18.

Hallin, C., Hofstede, J. L., Martinez, G., Jensen, J., Baron, N., Heimann, T., ... & Larson, M. (2021). A comparative study of the effects of the 1872 storm and coastal flood risk management in Denmark, Germany, and Sweden. *Water*, *13*(12), 1697.

IOC (2020). Quality Control of in situ Sea Level Observations: A Review and Progress towards Automated Quality Control, Vol. 1. Paris, UNESCO. IOC Manuals and Guides No.83. (IOC/2020/MG/ 83Vol.1) (2020). https://unesdoc.unesco.org/ark:/48223/pf0000373566.

Kirezci, E., Young, I. R., Ranasinghe, R., Muis, S., Nicholls, R. J., Lincke, D. and Hinkel, J. (2020). Projections of global-scale extreme sea levels and resulting episodic coastal flooding over the 21st Century. Scientific Reports 10(1, 11629) (http://www.nature.com/articles/s41598-020-67736-6) accessed December 26, 2022.

Melet, A., Irazoqui, M., Fernandez-Montblanc, T., Ciavola, P. (2021). Report on the calibration and validation of hindcasts and forecasts of total water level along European coasts, Deliverable 4.1 - ECFAS project (GA 101004211), www.ecfas.eu.

Nerem, R. S., Beckley, B. D., Fasullo, J. T., Hamlington, B. D., Masters, D., & Mitchum, G. T. (2018). Climate-change–driven accelerated sea-level rise detected in the altimeter era. *Proceedings of the national academy of sciences*, *115*(9), 2022-2025.

Pérez Gómez, B., Álvarez Fanjul, E., Pérez, S., de Alfonso, M., & 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.

Pérez-Gómez, B., García-León, M., García-Valdecasas, J., Clementi, E., Mösso Aranda, C., Pérez-Rubio, S., ... & Álvarez Fanjul, E. (2021). Understanding sea level processes during Western Mediterranean storm Gloria. *Frontiers in Marine Science*, *8*, 647437.

Pérez Gómez, B., Vilibić, I., Šepić, J., Međugorac, I., Ličer, M., Testut, L., ... & Zodiatis, G. (2022). Coastal sea level monitoring in the Mediterranean and Black seas. *Ocean Science*, *18*(4), 997-1053.

Pérez Gómez, B. and Lin-Ye, J. (2022). Quality information document. CMEMS-INS-QUID-013-053.

Product User Manual. For SEA LEVEL Reprocessed In Situ Product. INSITU_GLO_PHY_SSH_DISCRETE_MY_013_053. Pérez Gómez, B., Lin-Ye, J., Gallardo, A. and In Situ TAC partners. 2022.

Puertos del Estado (2019). SELENE: SEa LEvel NEar-real time quality control processing. Design & user's guide. version 1.0

Pugh, D. T. (1987). Tides, surges and mean sea level.Quality Information Document. IN SITU TAC. INSITU_GLO_PHY_SSH_DISCRETE_MY_013_053. Pérez Gómez, B., Lin-Ye, J., Bradshaw, E., Hibbert, A. 10-10-2022. Issue 1.0.

https://doi.org/10.5194/egusphere-2023-733
Preprint. Discussion started: 20 April 2023

Vousdoukas, M. I., Mentaschi, L., Voukouvalas, E., Verlaan, M., Jevrejeva, S., Jackson, L. P. and Feyen, L. (2018). Global probabilistic projections of extreme sea levels show intensification of coastal flood hazard. Nature Communications 9(1), pp. 2360 (http://www.nature.com/articles/s41467-018-04692-w) accessed December 26, 2022.

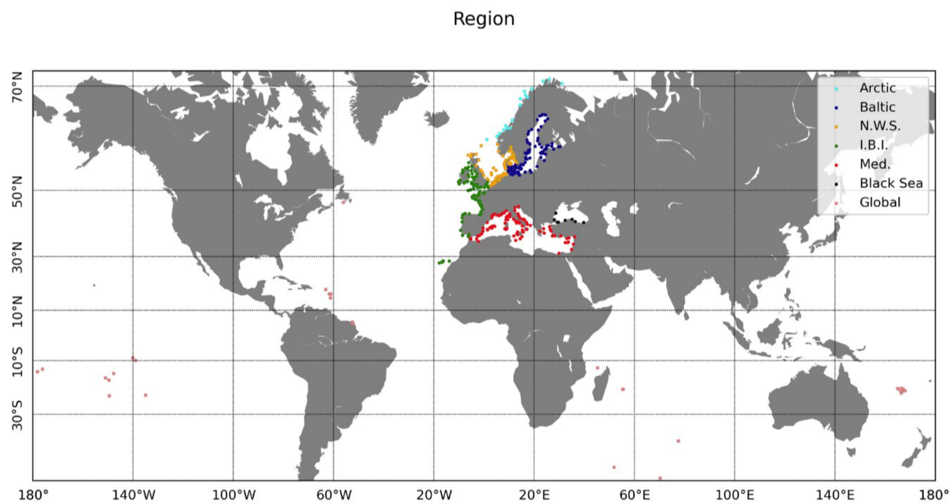

Fig. 1 Network of tide gauges in the Copernicus Marine Service In Situ TAC showing data integration by each of the regions (colours). Source: Quality Information Document, Issue 1.0.

https://doi.org/10.5194/egusphere-2023-733
Preprint. Discussion started: 20 April 2023

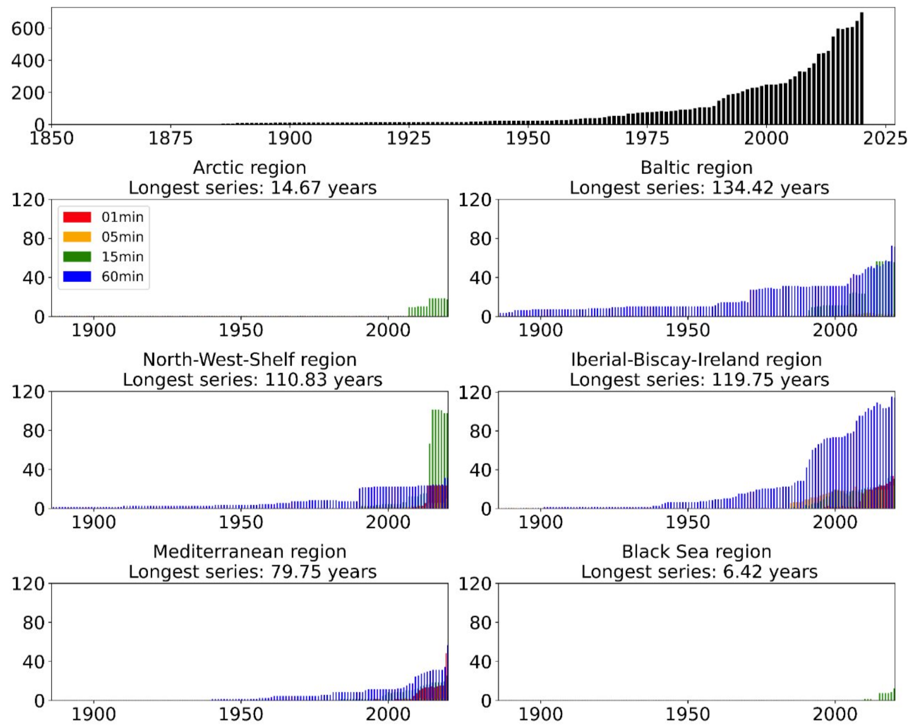

Fig. 2 Evolution of the number of tide gauges per year in Copernicus Marine Service In Situ TAC. It includes the series that span less than one year at the end of 2020. Product name: INSITU_GLO_PHY_SSH_DISCRETE_MY_013_053. Source: Quality Information Document, Issue 1.0.

https://doi.org/10.5194/egusphere-2023-733
Preprint. Discussion started: 20 April 2023

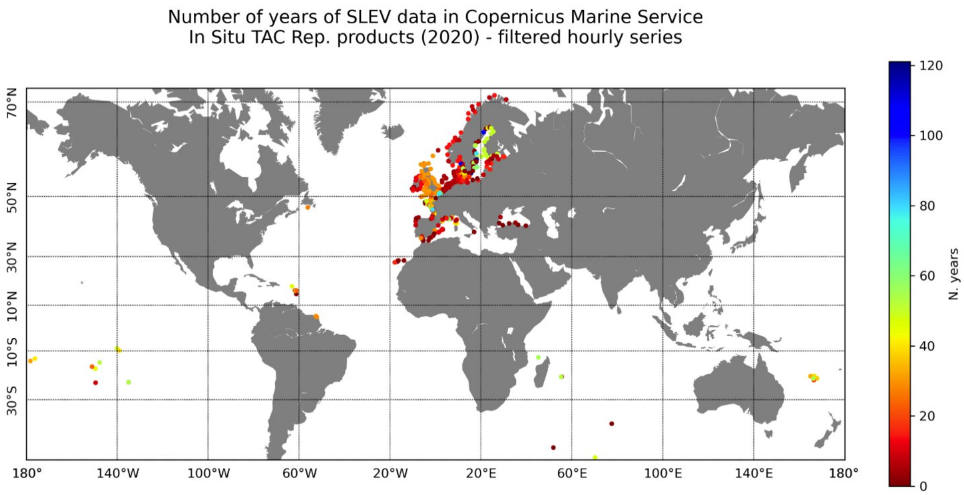

Fig. 3 Time series length of tide gauges that are included in the reprocessed product, first release. Source: Quality Information Document, Issue 1.0.

https://doi.org/10.5194/egusphere-2023-733
Preprint. Discussion started: 20 April 2023

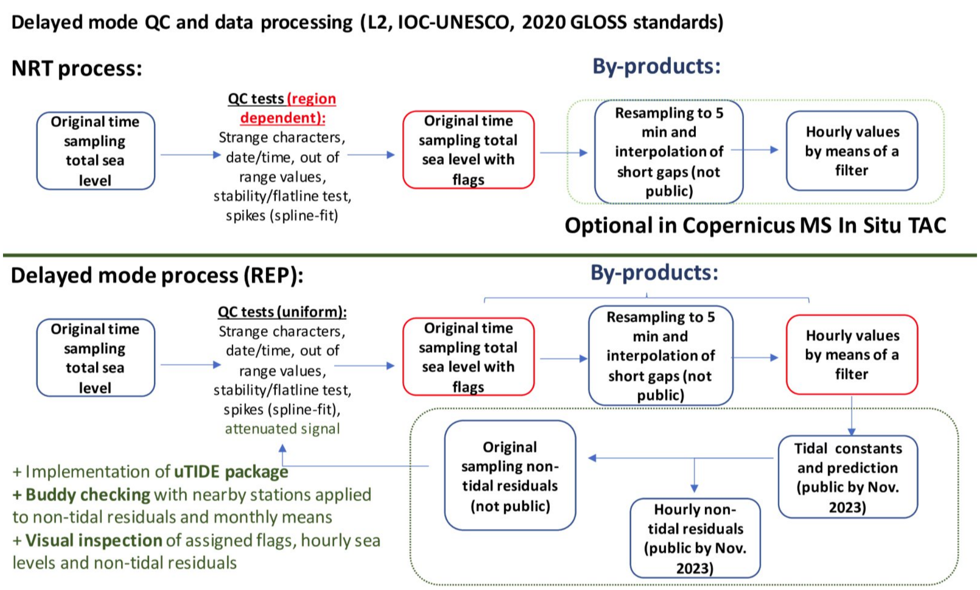

Fig. 4 Flow chart of NRT (near-real-time) process and delayed-mode reprocessing.

https://doi.org/10.5194/egusphere-2023-733
Preprint. Discussion started: 20 April 2023

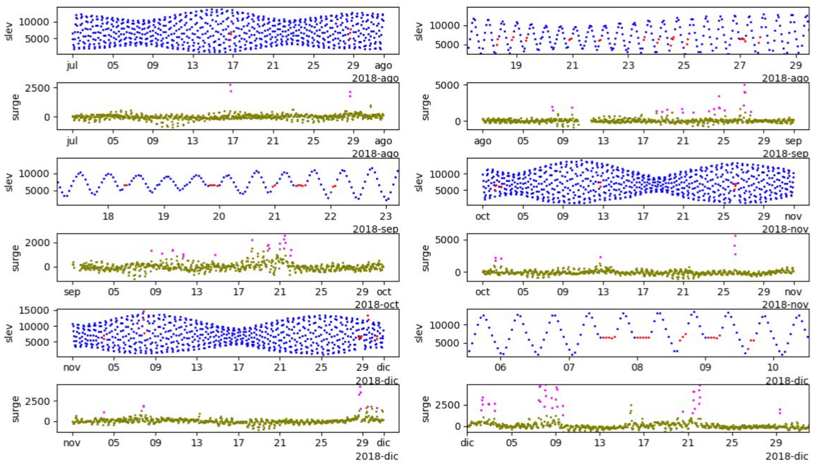

Fig. 5 Example of detection of bad data in the total sea level (slev), with the help of de-tided sea level (surge). Bad data are shown in red in the total sea level, and in pink in the de-tided sea level.

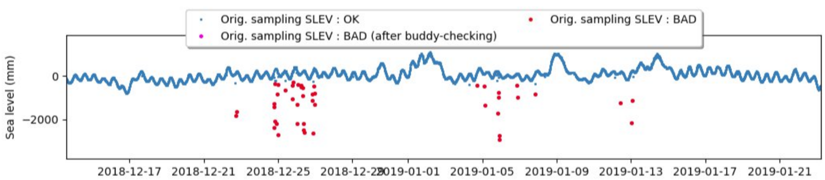

Fig. 6 Example of spike detection in the total sea level

https://doi.org/10.5194/egusphere-2023-733
Preprint. Discussion started: 20 April 2023

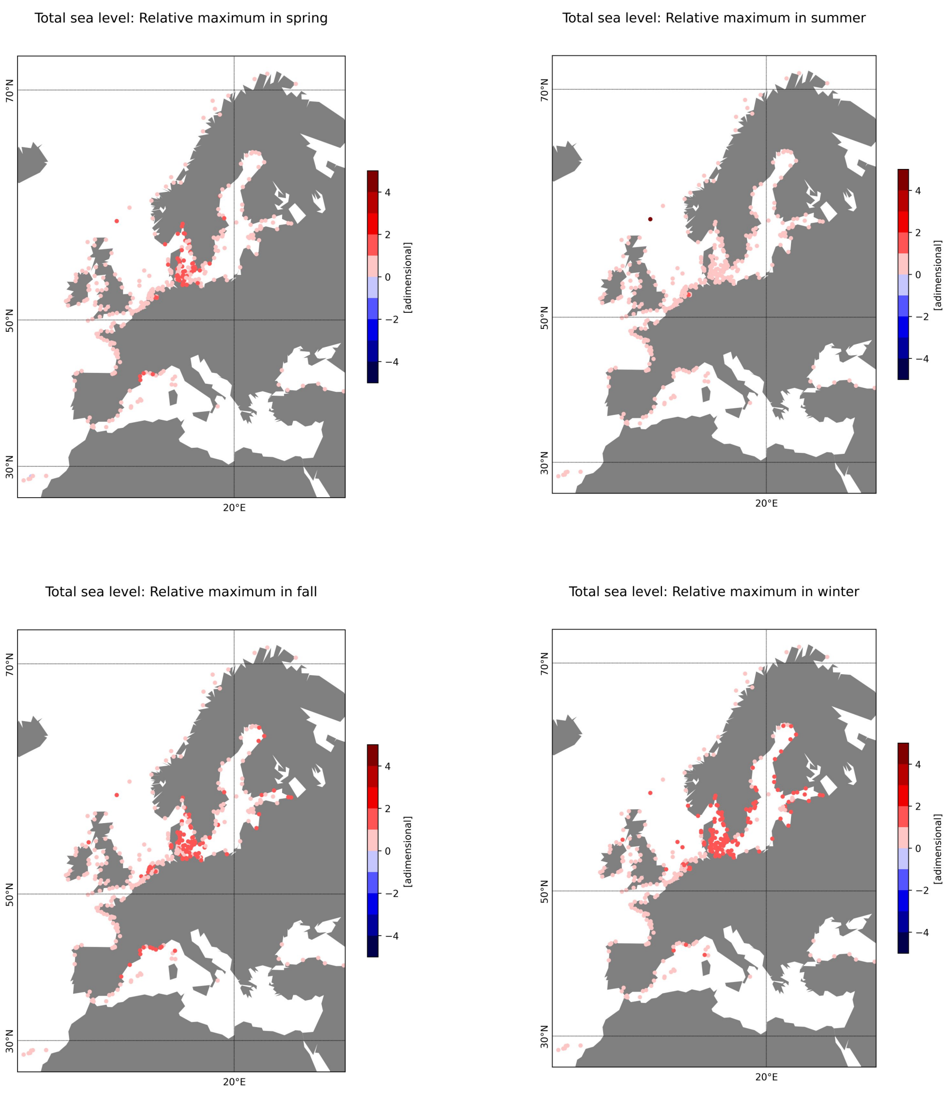

Fig. 7 Relative maximum sea levels, in each season.

https://doi.org/10.5194/egusphere-2023-733
Preprint. Discussion started: 20 April 2023

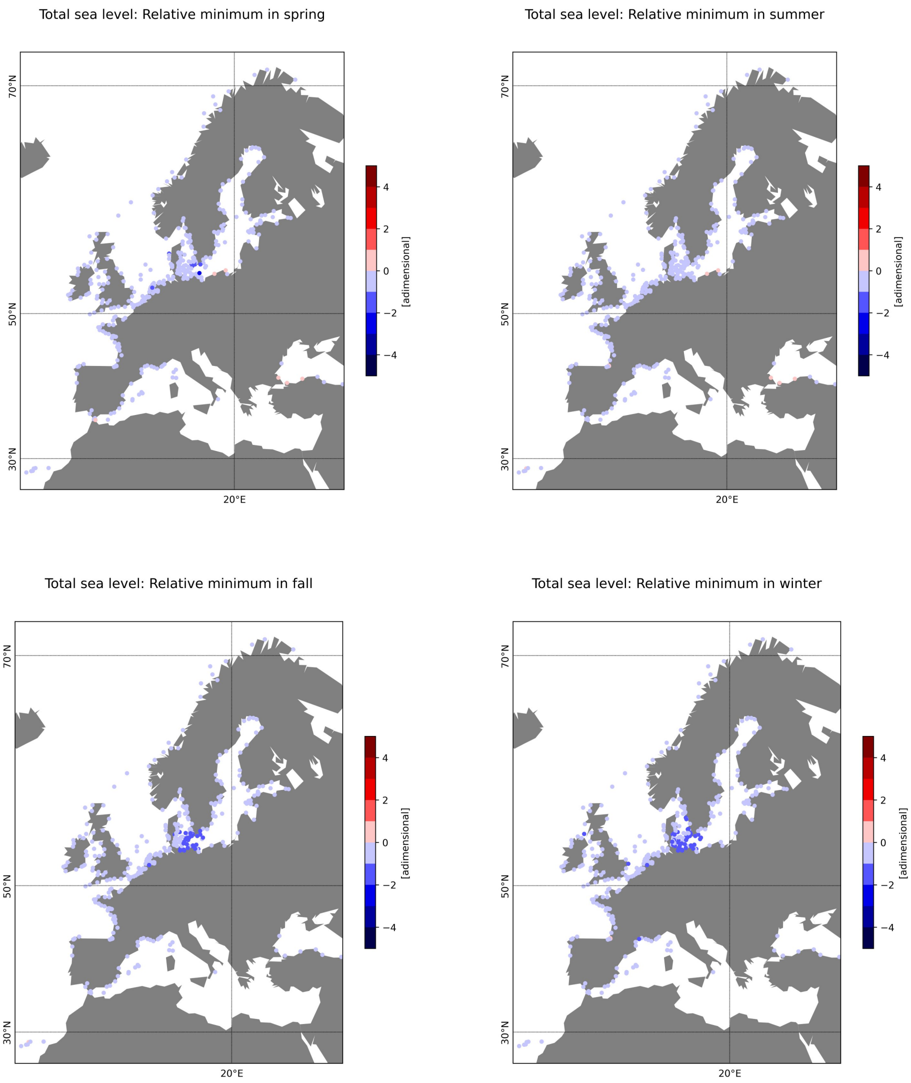

Fig. 8 Relative minimum sea levels, in each season.