IN SITU TAC

INSITU_GLO_PHY_SSH_DISCRETE_MY_013_053

Issue: 1.1

Contributors: Begoña Pérez Gómez, Jue Lin Ye, Angela Hibbert

Approval date by the Copernicus Marine Service product quality coordination team: 30/11/2023

CHANGE RECORD

When the quality of the products changes, the QuID is updated and a row is added to this table. The third column specifies which sections or sub-sections have been updated. The fourth column should mention the version of the product to which the change applies.

Issue	Date	§	Description of Change	Author	Validated By
1.0	28/08/2022	All	First version of the document	Begoña Pérez Gómez, Jue Lin Ye, Elizabeth Bradshaw, Angela Hibbert	Stéphane Tarot
1.1	10/08/2023	All	Second version: new dataset and improved quality tests	Begoña Pérez Gómez, Jue Lin Ye, Angela Hibbert	Stéphane Tarot

TABLE OF CONTENTS

I EXECUTIVE SUMMARY

I.1 I.1 Products covered by this document

This document provides a description of the procedures and tools used for the delayed mode quality control, processing and assessment of the Sea Level REProcessed (REP) product delivered by the Copernicus Marine Service In Situ Thematic Assembly Centre (In Situ TAC). It contains also a quality overview of the product, and applies to the following product as described in the Copernicus Marine Service Catalogue (Table 1):

Short Description	**Product code**	**Area**	**Delivery Time**
GLOBAL REP	INSITU_GLO_PHY_SSH_DISCRETE_MY_013_053	GLOBAL	Twice a year

Table 1: List of In Situ TAC products for which this document applies.

The Sea Level REP product is a global product that provides the best available version of in-situ historical sea level observations in In Situ TAC, reprocessed and validated in delayed mode at centralized level by the Production Centre. This is required for assessment of operational models and reanalysis, and for use by the research community. It integrates the in-situ sea level observations, mainly from tide gauges that provide coastal sea level, aggregated from the Regional EuroGOOS consortium (Arctic-ROOS, BOOS, NOOS, IBIROOS, MONGOOS and Black Sea GOOS), National Data Centers (NODC's), Hydrographic Offices or Metoffices, the Global Sea Level Observing System (GLOSS) and EMODnet Physics. The Copernicus Marine Service does not contribute to the maintenance or set up of the observing systems it uses. The complete list of variables distributed by the In Situ TAC can be found in the Copernicus Marine In Situ TAC physical parameters list (https://doi.org/10.13155/53381).

The product is updated twice a year, and validated with a delayed mode quality control and processing which includes visualization, computation of non-tidal residuals and hourly data, and check of datum stability by means of neighbour test or buddy-checking. The quality of the product is regularly assessed by means of specific metrics described in Section III.1 (Summary of metrics and tests used in the assessment).

Sea level data are available in the dedicated directory for sea level INSITU_GLO_PHY_SSH_DISCRETE_MY_013_053 of the Copernicus Marine Service Dissemination System. It includes three different datasets: the first one contains the diverse original time sampling data available at the In Situ TAC NRT, which vary from 1-min to 1 hour, with updated quality control flags; the second one contains the homogeneous hourly data timeseries, validated and filtered from higher frequency observations, when available; and the third one contains an hourly dataset including astronomical tide and surge/non-tidal residuals for the Iberian Biscay Ireland (IBI) region.

More detailed information can be obtained from the Copernicus Marine Service web page (https://marine.copernicus.eu/), the Copernicus Marine Service Desk (servicedesk.cmems@mercator-ocean.eu) and the In Situ TAC web page (http://www.marineinsitu.eu/).

I.2 I.2 Summary of the results

It is important to differentiate between the data validation process and the metrics employed for assessment of the product.

The data validation process is performed through automatic procedures including quality control tests, computation of by-products such as filtered hourly values, tide and non-tidal residuals and buddy checking or comparison with nearby stations. The latter two are useful to identify and flag less obvious problems in the original time series. These procedures, described in Section III.1 (Summary of metrics and tests used in the assessment), are complemented by a visual inspection of flagged data (to confirm the performance of automatic tests), sea level and non-tidal residual timeseries, and differences between nearby stations. Quality flags are set to inform the users of the level of confidence attached to the observations.

In addition, an overall assessment of the product is performed through the representation and computation of the spatial and temporal coverage, and an overview of the data quality (e.g., percentage of valid data) as described in Section IV (Validation results).

The main results of the validation process are the following:

- This product is based on in situ sea level data from 699 platforms, 9.4% more than the number of platforms available in the first release of the product, in November 2022. The difference is explained by the 60 new stations from Italy, Greece, Turkey, Cyprus, Israel and Egypt, improving significantly the spatial coverage in the Mediterranean Sea in this second release (November 2023).
- Most of these stations are in,European Seas with a still limited global coverage. Highest density of stations is observed in the Southern North Sea (Netherlands, Belgium, Denmark), although a high density can also be seen for other European countries (France, Spain, UK, Sweden, Italy). Fewer stations are present at higher latitudes, and practically none in the Southern Mediterranean Sea.
- The majority of these platforms are tide gauges installed in harbours, and therefore affected by local bathymetry and coastal processes.
- Brest (France), Marseille (France) and Kungsholmsfort (Sweden) tide gauges have the longest and continuous time series, starting in 1846, 1885 and 1886, respectively. All of them are still operational. Part of the oldest records have been recovered since the November 2022 release, improving the temporal coverage for several stations. St. Nazaire tide gauge, also in France, has the oldest recovered timeseries in the dataset, spanning from 1821 to 1925.
- The temporal coverage of sea level measurements start~~s~~ with a very low and stable number of stations since the end of the 19th century, increasing slowly until 1950, and increasing dramatically since then, especially over recent decades.
- Sea level data are provided with heterogeneous time samplings (e.g.: 1, 5, 10, 15, 60 min), depending on the data provider or main application, providing information of processes at different spatial and temporal scales. The longest time series have only hourly values, while modern stations provide higher frequency data, as required for extremes and tsunami warning. Hourly values are still the most common time sampling in In Situ TAC data files.

The percentage of good data in the whole dataset is very high (98.5%), as well as the average percentage of valid data per year (which we refer to as Completeness Index), which exceeds the 80% for most periods and regions, with very few exceptions (e.g. 36.7% in the Black Sea).

During the course of this reanalysis (2023), sea level observations available in In Situ TAC NRT Production Units have been validated and reprocessed. In future releases, the In Situ TAC will set up interfaces with new sources of in situ sea level observations such as GLOSS (Global Ocean Observing System) and EMODnet Physics. The product described here is focused on generation of a reprocessed and validated dataset of hourly sea levels, the most common time sampling available at the moment in the In Situ TAC. For those stations with only higher frequency data available, hourly values have therefore been obtained by means of a standard filter. A dataset with the original time sampling is provided, flagged during the delayed mode process. Original data providers may have additional higher frequency sampling data not distributed by the Copernicus Marine Service to date.

I.3 I.3 Estimated Accuracy Numbers

Table 2 summarizes the accuracy estimates of the sea level measurements that can be expected in the Sea Level REP product, depending on the time sampling. This is the best accuracy that a user can expect for the in situ data to which a 'good data' quality flag (see Table 4) has been applied after the validation process.

Variable	**Units**	**Accuracy of measurements**
Sea level (higher sampling or hourly instantaneous values)	cm	< 1
Sea level (filtered hourly values)	cm	< 1

Table 2. Accuracy numbers for sea level time series for different time samplings.

II PRODUCTION SYSTEM DESCRIPTION

Production Unit name: Puertos del Estado, Spain; NOLOGIN CONSULTING, Spain

Production system name: Global Ocean In Situ Reprocessed Sea Level (Copernicus Marine Service names: INSITU_GLO_PHY_SSH_DISCRETE_MY_013_053).

Production system name: Global Ocean In Situ Reprocessed Sea Level.

Description

The In Situ TAC is a distributed centre organized around 7 oceanographic regions: the global ocean and the 6 EuroGOOS regional alliances, 5 of them in Europe (ROOS: Regional Operational Oceanographic Systems) (see Figure 1). It involves 17 partners from 11 countries in Europe. It doesn't deploy any observing system and relies on data that are obtained exclusively funded by other sources than Copernicus Marine Service.

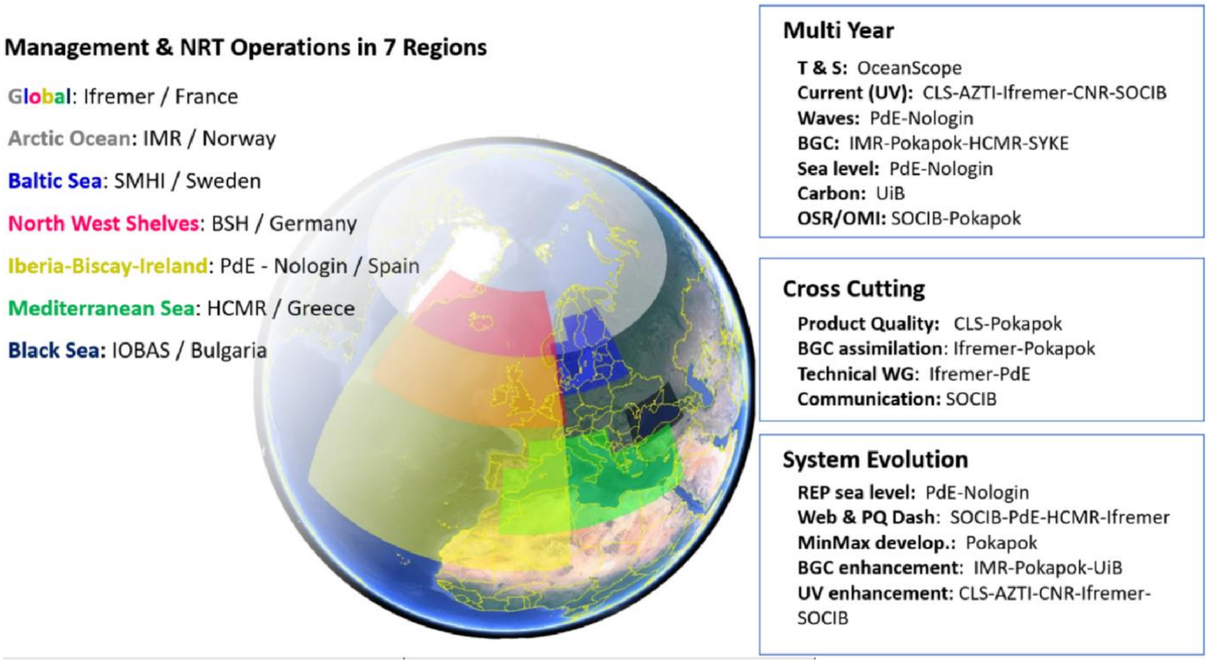


Figure 1: The Copernicus In Situ TAC organization and components – Leader: Ifremer/France

The In Situ TAC architecture is decentralized for the NRT production, but the production of the reprocessed multiyear products is centralized in a Production Centre composed by one or several institutions in charge of it. Figure 2 shows a scheme of the multi-year product operations in the In Situ TAC: national data providers feed marine data to Copernicus Marine Service through the regional alliances (global and regional: ROOSes) and international programs. The data are reprocessed in delayed mode and stored in the Production Centre (or Production Unit: PU) data base, starting from a frozen copy of NRT data. The reprocessed product is converted to a homogeneous NetCDF format and distributed, along with related Ocean Marine Indicators (OMIs) for some products, through the Central Dissemination Unit.

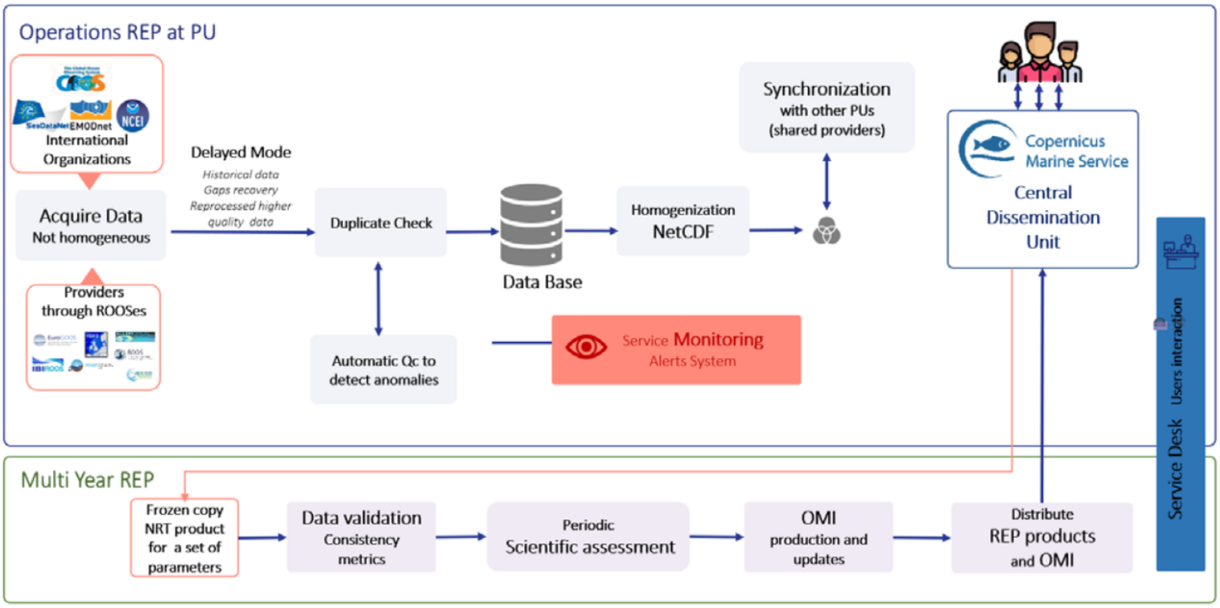


Figure 2: Scheme of multi-year product operations in the In Situ TAC. The reprocessing starts from a frozen copy of NRT data gathered from the regions and international programs; periodic validation and scientific assessment, conversion to NetCDF format, and computation of Ocean Marine Indicators are obtained and disseminated through the Central Dissemination Unit.

The In Situ Sea Level REProcessed product starts with a frozen copy of history files from the Global multiparameter Near Real Time (NRT) product (INSITU_GLO_PHYBGCWAV_DISCRETE_MYNRT_013_030).

The production consists of:

i) extraction of sea level data and metadata in the mentioned product (mainly from tide gauges);
ii) validation and generation of datasets with procedures detailed in this QuID, following best practices at global and European level (GLOSS, EuroGOOS DATAMEQ and EuroGOOS Tide GaugeTask Team);
iii) NetCDF files generation and check with a format and content checker;
iv) assessment of the product with the metrics described in the QuID;
v) preparation of documentation and distribution.

The whole validation and update process for this product is performed twice a year: around June for temporal extension of six months and at the end of the year (November-December) for temporal extension of six months and also several possible modifications or improvements of the product including full reprocessing.

III VALIDATION FRAMEWORK

The In Situ TAC is dedicated to assuring the accuracy of in situ observations through mainly two validation approaches: automatic tests applied in the Real Time Quality Control (RTQC) and the validation in delayed mode and assessment of the product.

For the first approach, a set of metrics were developed and described in detail in Copernicus for Near Real Time Quality Control and Validation of in situ sea level data (Pérez Gómez et al., 2010) and in IOC/UNESCO manual (IOC, 2020). These include automatic detection and flagging of:

1. invalid characters;
2. wrong date/time assignment;
3. out of range values (from predefined thresholds station dependent);
4. spikes detection (by means of a spline-fit algorithm); ~~and~~
5. constant values detection (stuck value or stability test).

These types of errors are easily detected and flagged on the original time series in NRT by means of existing automatic procedures, which can include the following L1 (Level 1 QC) by-products:

a) time series of homogeneous time sampling (interpolation of short gaps, flagged data and resampling);
b) non-tidal residuals computation;
c) hourly data obtained by means of an adequate/standard filter.

Computation of non-tidal residuals and hourly data is discretionary for In Situ TAC NRT data, and may be computed in different ways, depending on the region/data provider. To date (November 2023), none of the regions compute the non-tidal residuals in near-real time, despite their use in the automatic procedures of quality control improving significantly the detection of incorrect data. Some errors are easier to identify in the non-tidal residuals and in other by-products. Therefore it is possible to analyse again the original time series and flag and even correct the data, in an iterative process.

In order to improve consistency and homogenization of the Sea Level REProcessed product, a delayed mode quality control is applied to historical time series in the In Situ TAC NRT (L2 quality control according to UNESCO/IOC, 2020), including generation and quality control of non-tidal residuals, and filtered hourly data generation (when higher sampling data are available, see figure 3). As the product starts from a frozen copy of the history files in the In Situ TAC NRT, the steps of the quality control and data processing vary slightly depending on the original time sampling of the data, which are updated with new flags and filtered to hourly values, if these are not available (figure 3). Since the November 2023 release~~,~~ the tidal and non-tidal residual hourly components of sea level are distributed and included in a new dataset for the IBI region.

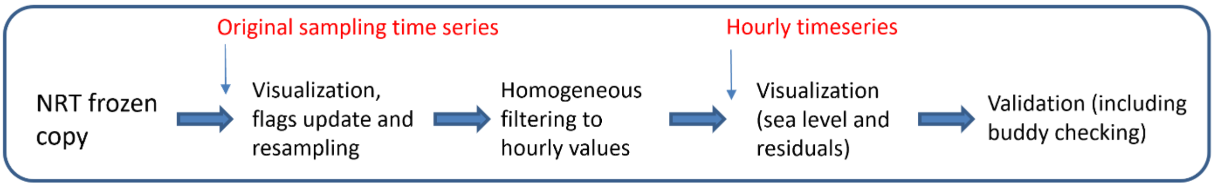


Figure 3: Scheme of multi-year sea level product operations in the In Situ TAC

Additional automatic procedures applied during the delayed mode validation of historical sea level observations include: attenuated signal detection, applied to hourly values, improved spikes detection tests, and "buddy-checking" applied to non-tidal residuals and monthly mean sea levels. The latter includes new automatic tests to identify possible datum changes. The whole process is complemented with a visual inspection of: original time sampling data, for validation of automatically flagged data, hourly sea levels and non-tidal residuals, and differences with nearby stations (buddy checking).

The complete list of metrics applied to sea level data (original time sampling and by-products) are presented in Table 3 and the QC flags assigned to data, according to In Situ TAC requirements, are shown in Table 4.

Short description of metrics	**Applicability of metrics**				
	Raw data	**By-products computation:**			
	Original time sampling	**Interpolation module (Resampled TS)** ✓	**Tide-Surge Module (Non-tidal residuals)** ✓	**Filter Module:Hourly data** ✓	**Monthly means** ✓
Strange characters detection	*X*				
Date and Time	*X*				
Out of range values	*X*		*X*		
Stability/flatline test	*X*		*X*	*X*	
Attenuated signal				*X*	
Spike detection	*X*		*X*		
Buddy-checking			*X*		*X*
Clock malfunction			*X*		
Datum change			*X*		*X*

Table 3: Metrics used for delayed mode quality control of the Sea Level REP product, according to GLOSS best practices (UNESCO/IOC, 2020). Different metrics are applied to the different timeseries (raw data and by-products), depending on the type of error. All by-products are computed during the re-processing in November 2023 (green check symbols). No metrics applied to resampled time series, intermediate by-product between the raw data and hourly values.

Code	Meaning	Comment
0	*No QC was performed*	-
1	*Good data*	All QC tests passed.
2	*Probably good data*	These data should be used with caution.
3	*Bad data that are potentially correctable*	These data are not to be used without scientific correction.
4	*Bad data*	Data have failed one or more of the tests.
5	*Value changed*	Data may be recovered after transmission error.
6	*Value below detection*	The level of the measured phenomenon was too small to be quantified/detected by the technique employed to measure it. The accompanying value is the quantification/detection limit for the technique or zero if that value is unknown.
7	*Nominal value*	Data were not observed but reported. Example: an instrument target depth.
8	*Interpolated value*	Missing data may be interpolated from neighbouring data in space or time.
9	*Missing value*	The value is missing, is not reported, is not applicable.

Table 4. Quality control flag scale in In Situ TAC, depending on the quality or a specific characteristic of each individual value (e.g., 1 for good data, if the value passed all QC tests, or 4 for bad data, if the value failed one or more tests

In order to monitor the quality of the product, several metrics are used to assess its spatial and temporal coverage, its evolution through the years and the quality status of the data.

III.1 III.1 Summary of metrics and tests used in the assessment

This section describes the procedures and metrics applied for the assessment and validation of the sea level REP or multi-year (MY) product. The assessment performed by the providers is not described in this document because it is different for each platform and variable in time, and this information is not generally provided in the metadata. A first survey was launched in 2022, in order to check which tide gauges in In Situ TAC had already been reprocessed by original data providers. It is important to emphasize that while some of these data providers may not perform (regular) delayed mode quality control and validation, others may have validated timeseries that have not yet been made available to In Situ TAC history repositories.

The situation is therefore heterogeneous and some countries are in the process of modernizing/updating their data distribution systems. Based on the input received from the different Production Units, delayed mode reprocessed data are propagated to In Situ TAC NRT history files for 6 institutions:

- UK National Oceanographic Centre: original sampling validated data until November 2019
- Puertos del Estado (Spain): hourly reprocessed and validated data updated until December 2022 for stations in the REDMAR network
- Marine Institute of Ireland: original sampling validated data until November 2019.
- Norwegian Hydrographic Service: updated the time series between 2007 and 2019, in 2020.

- Swedish Meteorological and Hydrological Institute: delayed mode reprocessed data made available to In Situ TAC both monthly and annually.
- SHOM (France): files of 10-min and hourly values have been validated by SHOM for those tide gauges belonging to the RONIM network, until March 2021.
- The Mediterranean NRT Production Unit applied quality control including filter to hourly values to Mediterranean stations.

A more detailed description of the data processing, validation metrics and related automatic procedures used for generation of the product (see Table 3) is provided in Sections III.1.1 and III.1.2. . In 2023, two new automatic tests were implemented. Therefore, the product was not only extended to December 2022, but also reprocessed for the whole timeseries.

III.1.1 Tide computation and by-products generation

Computation of sea level by-products, derived from original data, are an essential component of the validation process. The following data processing and by-products generation are performed automatically:

Interpolation and resampling:

Most of the raw data from a tide gauge arrives with a data sampling of several minutes (1, 5, 6, 10, etc), dependent on the data provider. Sometimes~~,~~ the data sampling is not even regular, or data arrive at unexpected time stamps (e.g.: 5 min data arriving at 02, 07, 12, instead of 00, 05, 10). By November 2023 a growing number of institutions provide 1 min or less sampling data for extremes/tsunami warning and monitoring. All this explains the large heterogeneity of raw data sampling found in the In Situ TAC NRT dataset.

As a previous step before data filtering to hourly values raw data are interpolated and resampled by checking and adjusting the time interval; by interpolation of isolated spikes previously flagged by the quality control; by filling gaps with correct date assignment and special value for null-values; and by interpolation of very short gaps (10-25 min, configurable depending on the tidal range).

As a result new homogeneous and “clean” 5-min averaged timeseries are generated for all the stations, filtering higher frequency oscillations.

Filter to hourly values

Hourly values have been considered adequate for tide computation and operational oceanography for many years, becoming one of the standard products of sea level data. These must be obtained by means of a suitable filter, depending on the original data sampling. We use the 5-min sampling timeseries described above to obtain hourly values by means of the following symmetrical filter for 5-min data in eq.1, as described in Pugh (1987):

$$slev_f(t) = F_0 \cdot slev_5(t) + \sum_{m=1}^{M} F_m[slev_5(t+m) + slev_5(t-m)] \tag{1}$$

where M=54, $slev_f(t)$ is the hourly filtered value at time t, $slev_5$ are the sea level values for 5 min data, and $F_{0 \ldots m}$ the weights applied to the 5-min data (IOC/UNESCO, 2020, Pugh, 1987).

Tide and non-tidal residuals computation

Harmonic analysis of hourly sea level values is carried out using uTide, a software developed by the University of Rhode Island and adapted to Python (Codiga, 2011). Harmonic constants (amplitude and phase of the astronomical tidal constituents) are obtained on a yearly basis (67, for a complete year), and used for computation of the tide and the detided time series (non-tidal residuals) of each year. Only those years with large gaps (>0.5 week) adding up to less than 1 month are considered. In addition, the ~~following~~ condition in eq.2 must apply for all months:

$$log_{10}\frac{max(A)}{min(A)} \leq 0.7 \qquad (2)$$

where A is the monthly amplitude of M2 (main semidiurnal tidal constituent). If a year does not comply with these conditions and astronomical constants are not computed for that year, the de-tiding process can borrow the harmonic constants from a more complete year of data, selected from up to three years before or after that year for its computation.

As described in UNESCO/IOC, 2020, the de-tided sea level can help pinpoint anomalous behaviour. 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.

De-tiding and non-tidal residuals computation allow the whole quality control process to be applied to both the total and detided sea level signal, as recommended by UNESCO/IOC, 2020. For the new dataset delivered for the IBI region in November 2023, the tidal and non-tidal residual components are also distributed. For this dataset, annual harmonic constants and tide predictions are obtained from hourly data as described in this section, including the whole set of 67 harmonic constants. The non-tidal residuals are then computed as the difference between hourly observed and predicted values, and a low-pass filter (cutoff period=15 h) is applied to deal with the tidal remnant from harmonic analysis based on hourly values at some stations (Horsburgh and Wilson, 2007; Pérez González et al., 2017).

III.1.2 Automatic tests applied in delayed mode

Most of the checks applied in delayed mode are those performed in real time, which are repeated for a better detection of incorrect data (e.g. spikes). Delayed mode tests ensure all the checked values are in the center of a moving window, something which is not possible when applying the tests to the last value received in real time.

Out of range values detection:

This test allows detecting and flagging values exceeding predefined upper/lower limits for each station. Sea level presents a large spatial variability, mainly due to the spatial variation of the tidal range: from 16 m in the Bay of Fundy, Canada, the highest tidal range in the world, or 14 m in the Mont Saint-Michel, France, the highest range in Europe, to practically zero at many locations in the Mediterranean Sea. This range may differ within a region depending on the characteristics of the site (higher values found at the inner parts of an estuary), or due to the chosen datum or reference of the measurements.

To account for this variability, upper and lower limits have been defined for each station from the 99th percentile ($q99$) of historical data, as $\pm 3\ q99$. The same procedure is applied to non-tidal residuals.

Stability (flat line) test:

A flag is assigned when there is no change in the magnitude of sea level after a specified number of time steps (stuck values). The number of data values or time steps to begin to flag ($limit_{stuck}$) may be different depending on the time sampling of the data, the sea level units (mm or cm) and the natural sea level variability. For this product, the criterium expressed by eq.3 has been applied, based on some of these characteristics:

$$limit_{stuck} = 10 + C_{stuck,precision} + C_{stuck,series} + C_{stuck,visual} \tag{3}$$

Where the values of C_{stuck} are defined, based on the experience of processing timeseries with different types of sea level variability and quality, in the following way: $C_{stuck,precision}$ is set to 15 if the precision is in cm, and to 0 otherwise; $C_{stuck,series}$ is set to 10 for those stations where over 50% of the historical series of total sea level has stuck data (same value repeated for at least two time intervals), and is set to 0 otherwise; $C_{stuck,visual}$ is 0, a priori, and may be modified when, after a visual inspection of the output of the test, many false stuck values are observed. This may happen in stations with very small sea level signal and precision (cm). In this case, the test is run again with values of $C_{stuck,visual}$ becoming 20, for time samplings ≥ 10 min, and 40 for time samplings < 10 min. This process can be repeated, if necessary, to avoid false flagging of possible correct data.

Attenuated signal:

This procedure allows detection and flagging when the signal is not a flat line, but diminishes over a number of cycles. It is applied to hourly values, where this fault is more readily detectable than in higher frequency data. The test is applied to a 12-h moving window centered at the value *i*, and its variability is checked with respect to the 75th percentile (q_{75}) and the standard deviation std, of the historical time series. Data are flagged when the conditions in eq.4 and 5 are met:

$$slev(i) - \overline{slev}(i) < q_{75} / C_{att,1} \tag{4}$$

$$std(i) < std/C_{att,2} \tag{5}$$

where $slev(i)$ is the i^{th} hourly sea level, $\overline{slev}\ (i)$ is the i^{th} rolling mean sea level, $std(i)$ is the i^{th} rolling standard deviation, and $C_{att,1}$,$C_{att,2}$ are constants dependent on the tidal range or sea level variability of the station, or even on the tidal phase (spring/neap tides) in many cases. As an example, for macro-tidal ranges and spring tides ($q_{75} > 1.0\ m$ over mean sea level), the test works well with $C_{att,1} = 2$ and $C_{att,2} = 3$.

Spikes detection:

A local spline-fit algorithm is applied to a moving window of several hours, to detect spikes, discarding those data exceeding n-sigma from the fit. This algorithm is applied to the original sampled data and non-tidal residuals (less than 1 hour sampling), with the following characteristics and values by default:

- Total sea level: 2-degree polynomial fit, n-sigma=3, local window size:
    - Time sampling < 5min: 500 points
    - Time sampling ≥ 5 min and < 15 min: 100 points
    - Time sampling ≥ 15 min: 40 points
- Non-tidal residuals: 3-degree polynomial fit, n-sigma=5, local window size:

  - Time sampling < 5min: 1000 points
  - Time sampling ≥ 5 min and < 15 min: 200 points
  - Time sampling ≥ 15 min: 80 points

These values can be modified for individual stations depending on sea level variability or number of incorrect data.

A complement to the spline-fit spikes detection algorithm described above was implemented for the second version of the product in November 2023, in order to improve detection and flagging of persistent spikes or outliers. It is applied immediately after the previous one to original sampling data (total sea level and non-tidal residuals). The gradients described in eq.6 and 7 are computed for each timestep $t(i)$:

$$gradient = \frac{slev(i+1) - slev(i)}{t(i+1) - t(i)} \qquad (6)$$

$$gradient_{rev} = \frac{slev(i-1) - slev(i)}{t(i-1) - t(i)} \qquad (7)$$

when the module of the gradient is greater than a defined tolerance, and the difference between the $slev(i)$ and and the median of $slev$ is greater than twice the 99th percentile of $slev$, we start flagging the $slev(i)$ with QC=4. When the difference between $slev(i)$ and the median of $slev$ goes back to be smaller than thrice the 99th percentile of $slev$ and the module of $gradient_{rev}$ is smaller than the tolerance for gradient, we stop flagging $slev(i)$ with QC=4.

Buddy checking or “neighbour” test:

Apart from the quality control and inspection of non-tidal residuals, comparison with a second sensor at the same location, or with data from a nearby station, is a useful test for detection of less obvious problems in the data (e.g. datum changes and drifts, dubious extremes, etc). This test is applied to hourly non-tidal residuals (detided time series) and to monthly mean sea levels.

The algorithm selects two neighbouring stations for each target station, searching for those available in a radius from 0.1° to 0.5°. If no stations are found, this test is not applied. Otherwise, differences are computed between the values of the target station and the nearby ones. Data are then flagged in the following way:

- Detided time series: if differences exceed 1.0 m
- Monthly means: if differences exceed 0.5 m

This method is dependent on the availability and the quality of the nearby stations. An improved version of this test will be implemented in future releases with the use of altimetry and model data.

A first update of this test for monthly means was implemented for the November 2023 release. The gradient described in eq.8 is obtained for the monthly means of the target station and its neighbouring station:

$$gradient_{monthly} = \frac{slev_{monthly}(i+1) - slev_{monthly}(i)}{t(i+1) - t(i)} \qquad (8)$$

If the difference between $gradient_{monthly,target}$ and $gradient_{monthly,neighbor}$ is greater than 100 mm/month for over 3 months, then we start counting the period as having a potential datum problem, the period is automatically flagged, and then the flag confirmed later during the expert visual inspection of the data.

III.1.3 Visual inspection of the timeseries

Visual inspection is essential for validation of the flags assigned by the automatic tests, and one of the main advantages of the delayed mode validation process. Visual inspection of non-tidal residuals and differences with nearby stations is particularly important for detecting clock malfunction or datum changes, as these type of errors, very common in tide gauges timeseries, are difficult to detect or correct with existing automatic procedures.

Figure 4 provides an example of detection of a clock malfunction problem, thanks to visualization of non-tidal residuals, which show an oscillation during the effected period. Once detected visually, it is possible to go back to the original time series of total sea level and flag manually the period. A flag QC=3 is assigned (orange colour in Figure 4), corresponding to bad data that can be potentially corrected. For now, this type of error is not flagged automatically.

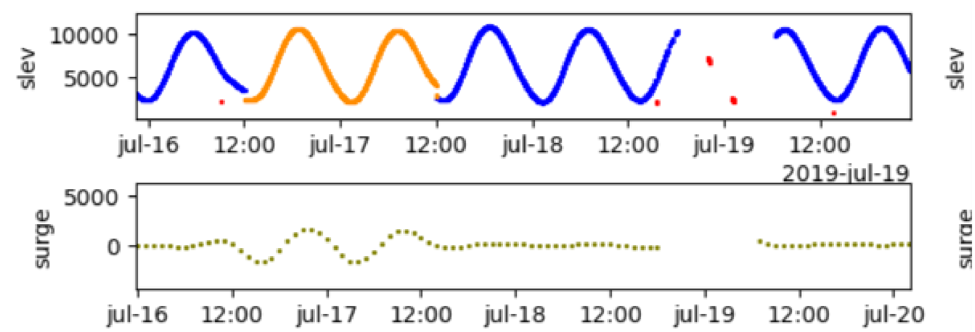


Figure 4: example of detection of a clock malfunction through visualization of the non-tidal residuals (bottom). A QC=3 has been manually assigned for the affected period in total sea level time series (top, data in orange colour), when oscillations are observed in the non-tidal residuals time series. Other automatically flagged data (QC=4, red colour in the top panel) can be seen in the total sea level plot.

Figure 5 presents another example where visualization helps to confirm the adequate flagging of automatic algorithms. In this case, the spline-fit algorithm applied to non-tidal residuals has correctly flagged data which had not been flagged by the test applied to total sea level time series. The algorithm has correctly propagated these flags (QC=4, red colour) to original time series of total sea level.

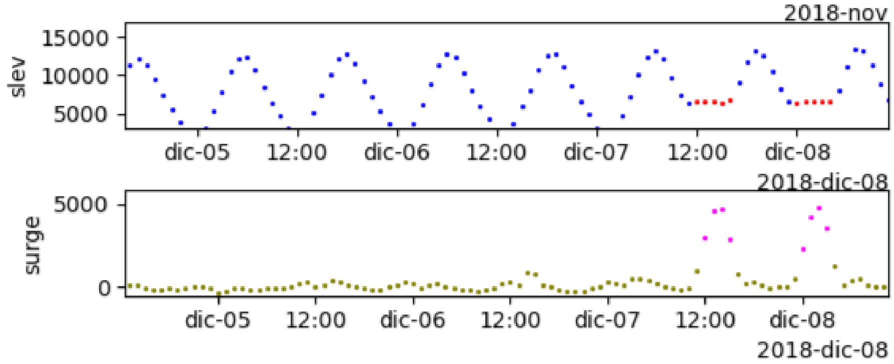


Figure 5: example of automatic detection of bad data thanks to application of tests applied to non-tidal residuals signal (bottom panel). A QC=4 is propagated to the total sea level time series (top, data in red colour).

IV VALIDATION RESULTS

This section presents the results of the metrics for assessment of the product in terms of temporal and spatial coverage for two datasets provided (original time sampling data with flags, and validated/reprocessed hourly sea levels).

IV.1 Temporal coverage of the Sea Level REP product

This section presents the temporal evolution of the number of platforms providing in situ sea level measurements in Copernicus Marine Service In Situ TAC. For now, these data come from tide gauges installed along the coastlines, most of them in harbours.

The total original number of platforms in November 2023 is 699 (Figure 6, top). The earliest timeseries corresponds to Saint Nazaire tide gauge, in the French Atlantic coast, which began in 1821, with intermittent data coverage until 1925. The longest and more continuous time series in this product corresponds to the Brest (France) tide gauge, starting in 1846 and still operational today (2023). Also in France (Mediterranean coast), the Marseille tide gauge time series started in 1885 and it is also active today in 2023 (138 years). In the Baltic Sea, four Swedish stations started operations in 1886, and one of them is still operational (Kungsholmsfort), becoming one of the longest time series in this product (136.5 years). The total number of stations is very small and practically constant until 1950. Since then, the number of stations increases very slowly until the mid 1980's, in the 20th century, after which the rate of increase becomes higher until 2004. The highest growth in the number of platforms is observed since 2004/2005 until 2022, where the largest value is reached, with no apparent decrease observed in recent years. This may be partially explained by the new tide gauges installed over the last two decades for tsunami warning, harbour operations and real time warning systems. Data start in the 2020's for the Arctic and the Black Sea regions.

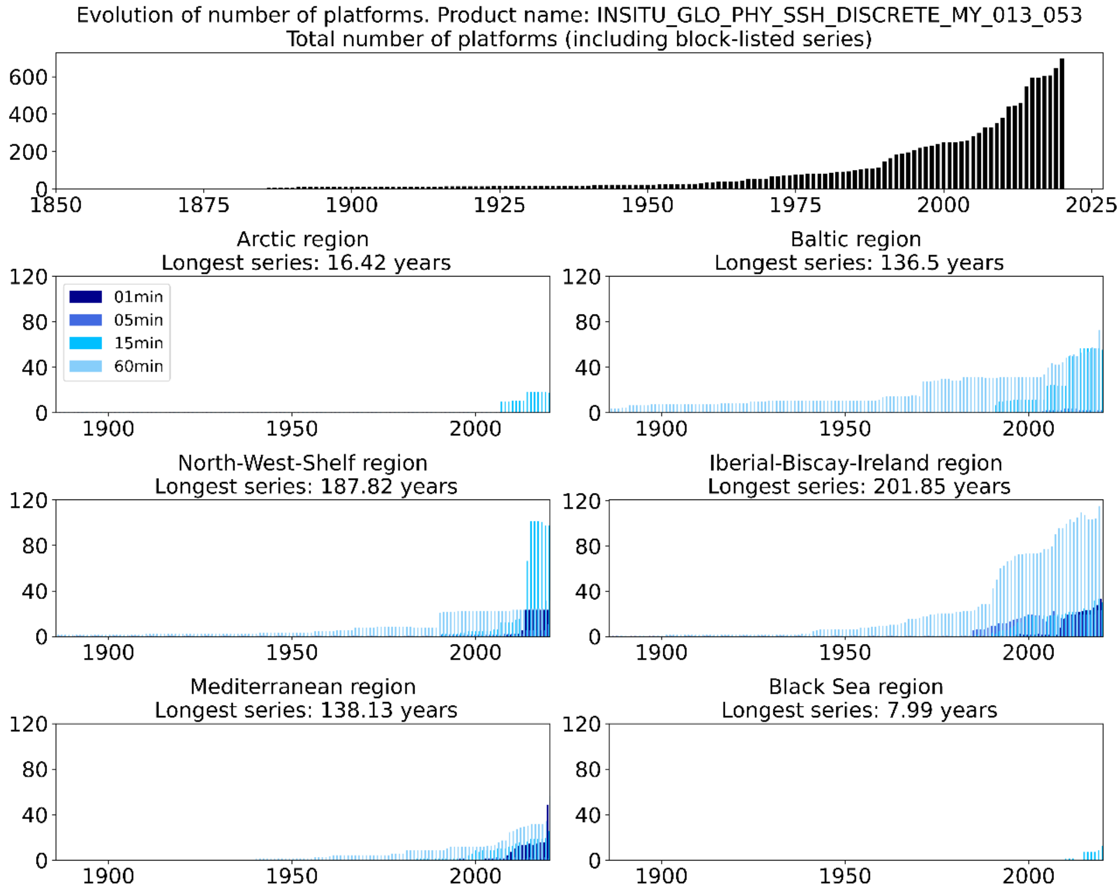


Figure 6: temporal evolution of the number of stations or platforms per year, from the sea level reprocessed product, for all the regions (top panel) and for different geographical regions (bottom panels), depending on the original time sampling data.

Original time samplings of In Situ TAC NRT history files are heterogeneous and dependent on the data provider and/or the regional Production Unit data distribution strategy. Figure 6 also shows, for the different regions, the number of stations providing data with different time intervals (original time sampling dataset) over the years. For some of the platforms, such as those operated by France, data are available at different sampling intervals. While most of the long historical records have only hourly values, common time intervals available in recent years are typically 1, 5, 10 and 15 min. Most of the stations in the Iberia Biscay Ireland area, the North Sea, the Baltic Sea and the Mediterranean Sea contribute with hourly data to the In Situ TAC MY Sea Level product, despite the large number of stations providing 1 min data in recent years.

Only shorter timeseries are available, starting this century, in the Arctic and the Black Sea regions, which contribute only 15 min data to this product.

For those tide gauges with only higher frequency data available, hourly values have been computed and filtered as described in Section III, and included in the dataset of validated and reprocessed hourly values. Table 5 shows a quantitative overview of the temporal coverage evolution in this dataset for the different regions, divided in three reference periods: 1880-1950, 1951-1985 and 1986-2022. The first one is the longest (70 years), corresponding to the oldest historical records, when less tide gauges were in operation and/or when less data are available in In Situ TAC. The other two are close to 35 years periods (35 and 37), the last one including the satellite altimetry era and the decades of larger increase on data availability. The metric includes the mean number of stations per year and the standard deviation within each period in total (global, first row) and for each geographical region.

Hourly sea levels dataset	**Number of platforms**					
	1880-1950		**1951-1985**		**1986-2022**	
	$\overline{N}$	**Std**	$\overline{N}$	**Std**	$\overline{N}$	**Std**
GL	**13.7**	**5.0**	**55.3**	**24.4**	**286.1**	**151.9**
AR					14.1	4.0
BO	7.9	1.9	19.4	8.4	65.5	34.5
BS					11.2	8.6
NWS	1.8	0.7	5.9	1.9	52.1	53.7
MED	1.0	0.0	4.7	2.4	46.1	32.4
IBI	3.8	2.0	16.2	5.3	84.5	27.8

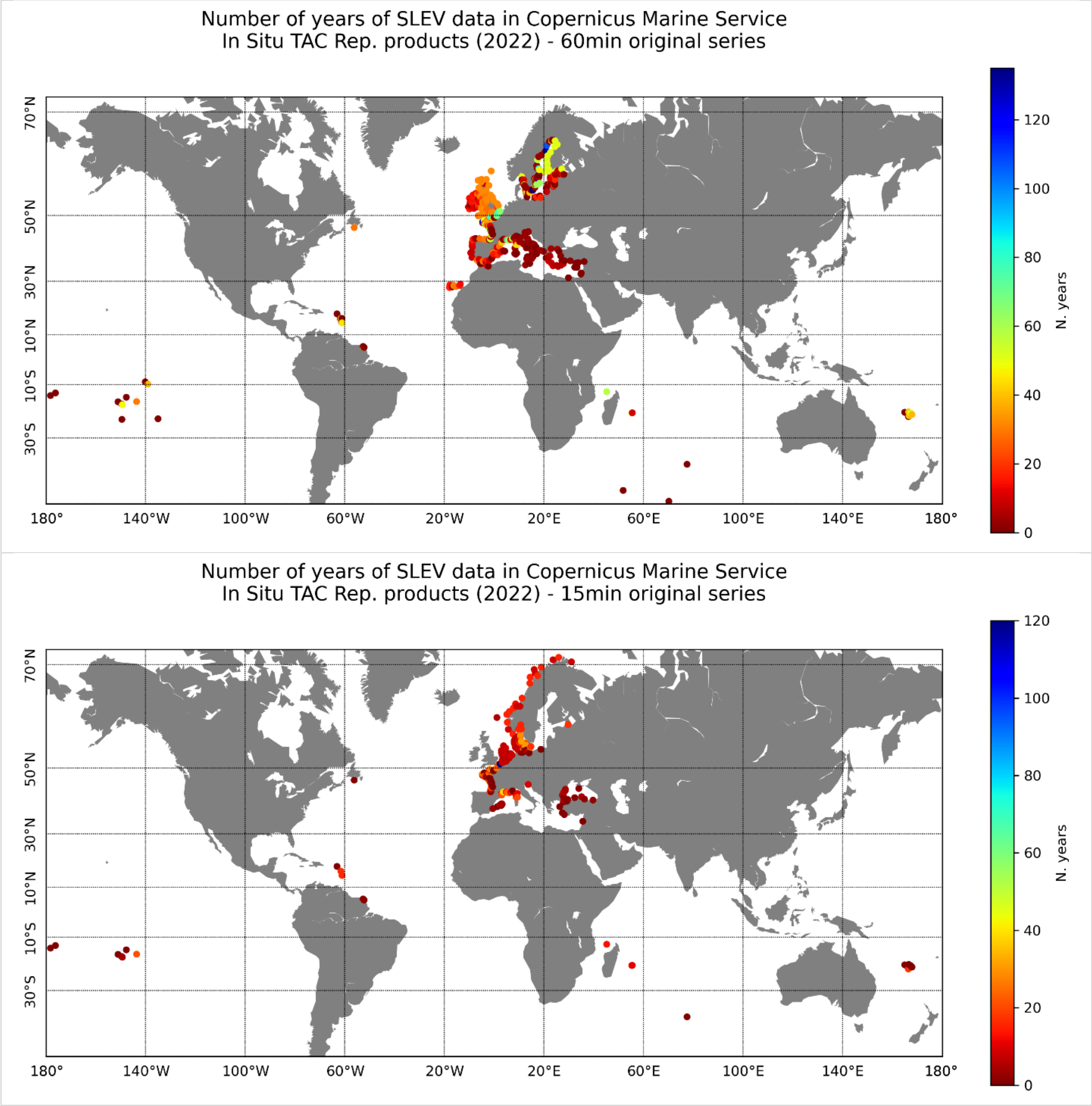
Table 5. Mean number of platforms ($\overline{N}$) and standard deviation (Std) for different regions and three periods: 1880-1950, 1951-1985 and 1986-2022. GL: global (all stations); AR: Arctic region, BO: Baltic Sea; BS: Black Sea; NWS: North-West-Shelf; MED: Mediterranean and IBI: Iberia-Biscay-Ireland region.

The total average number of stations $\overline{N}$ is obviously very low in the first period, with no data at all in the Arctic and the Black Seas and most of the platforms being in the Baltic Sea. The number of stations increases dramatically in the second period (for all regions except the Arctic and the Black Sea) and for the third period (all regions), almost multiplied by 5 in both cases (from 13 to 55 and to 286, respectively).

IV.2 Spatial coverage of the Sea Level REP product

The spatial coverage of the original time sampling dataset, colour scaled with the number of years, is shown in Figures 7 (60 and 15 min samplings) and 8 (5 and 1 min samplings). The number of stations and the length of the timeseries vary for the different time intervals. Some of the stations (e.g. French operated stations) have files with different sampling intervals (and different length). For the rest of networks, there is usually only one time sampling available: e.g.: hourly data from Spain, Portugal, UK, Ireland and Sweden, and 15 min data from Norway, Germany, Denmark, Bulgaria, Romania or Türkiye. The largest number of stations is found in hourly files, followed by the ones with 15 min data. These contain also the longest time series.

Most of the stations are located in Europe, except a few operated by France (SHOM) in the Caribbean, the South Pacific and the Indian Ocean, showing the lack of global coverage of the product at the moment. There is a high density of stations along most European coasts, especially in the Southern North Sea (Netherlands, Belgium, Denmark), but also along the coasts of France, Spain, UK, Sweden, Italy and Turkiye. The density is lower at higher latitudes and there is a gap in the Southern part of the Mediterranean Sea (North of Africa countries).



Number of years of SLEV data in Copernicus Marine Service
In Situ TAC Rep. products (2022) - 60min original series

Number of years of SLEV data in Copernicus Marine Service
In Situ TAC Rep. products (2022) - 15min original series

Figure 7: Stations in In Situ TAC with original sampling = 60 min (top) and with original sampling = 15 min (bottom). Colour scale is number of years.

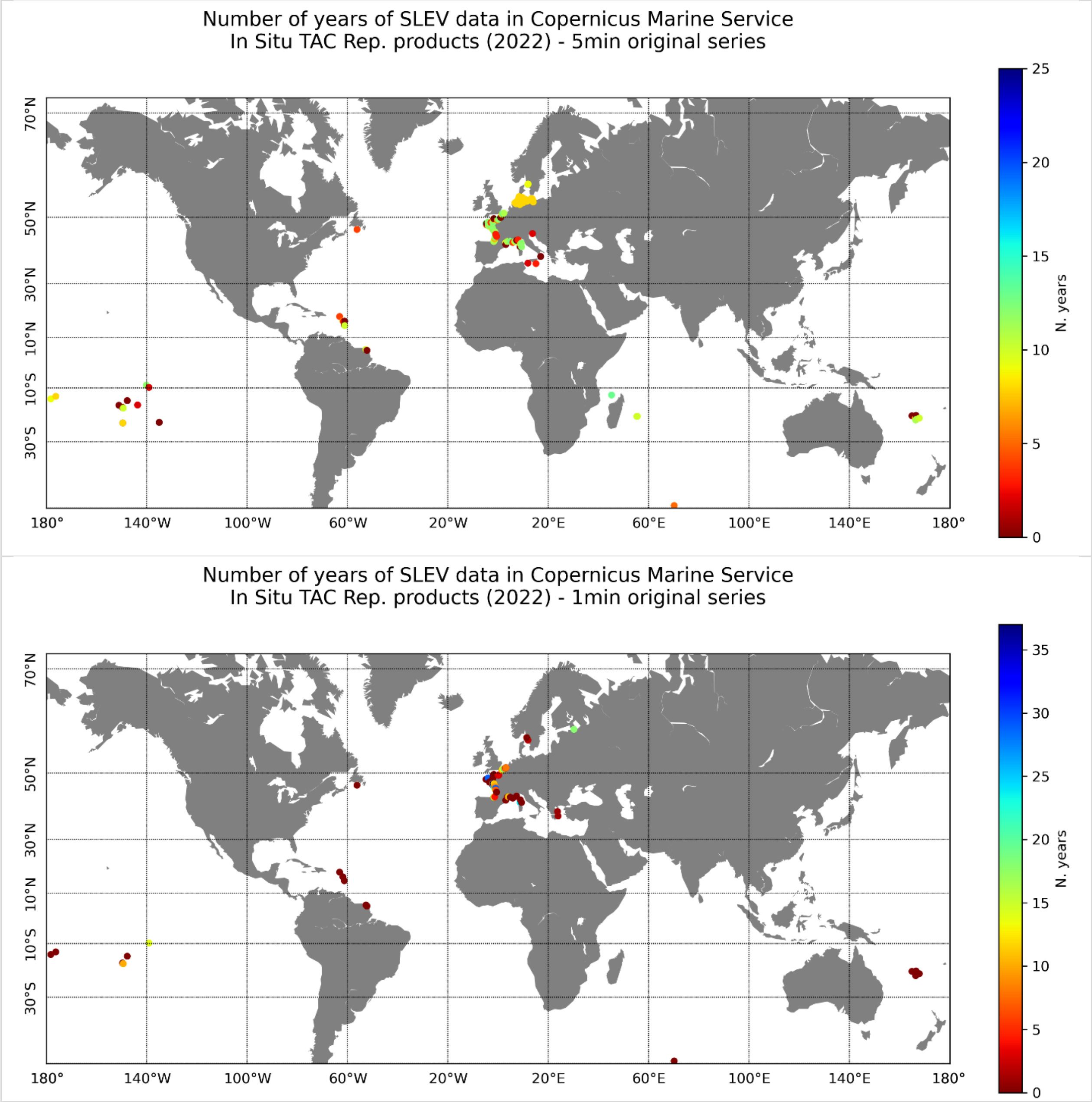


Figure 8: Stations in In Situ TAC with original sampling = 5 min (top) and with original sampling = 1 min (bottom). Colour scale is number of years

Figures 9 and 10 provide the spatial coverage of hourly sampling data in Europe, for the original sampling dataset (Figure 9), and for the final hourly values dataset created during reprocessing and delayed mode validation (Figure 10). Application of the automatic procedures including the filter to hourly values has allowed to increase significantly the number of stations with homogeneous hourly reprocessed values in In Situ TAC, for regions where these data were not previously available: tens of stations have now been added to the dataset in the North-West-Shelf region (especially in the south and around Denmark), and several ones appear now in the Arctic (Norway) and the Black Seas.

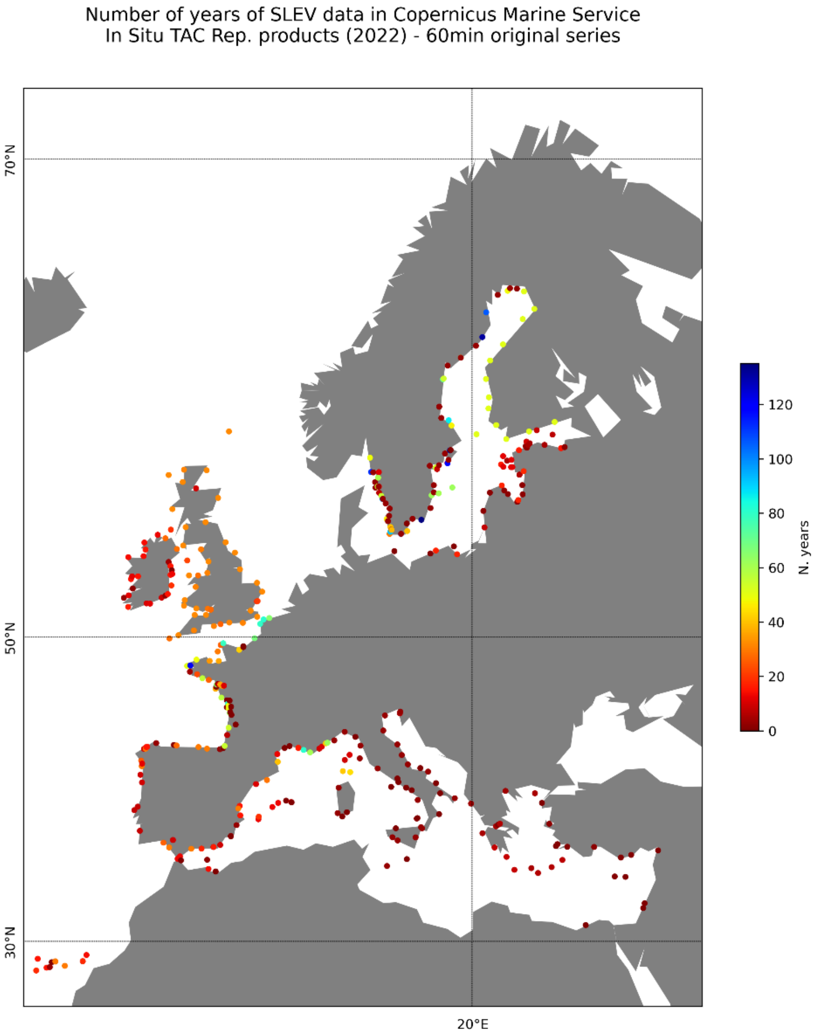


Figure 9: Stations with original sampling 60 min, European region (colour scale number of years).

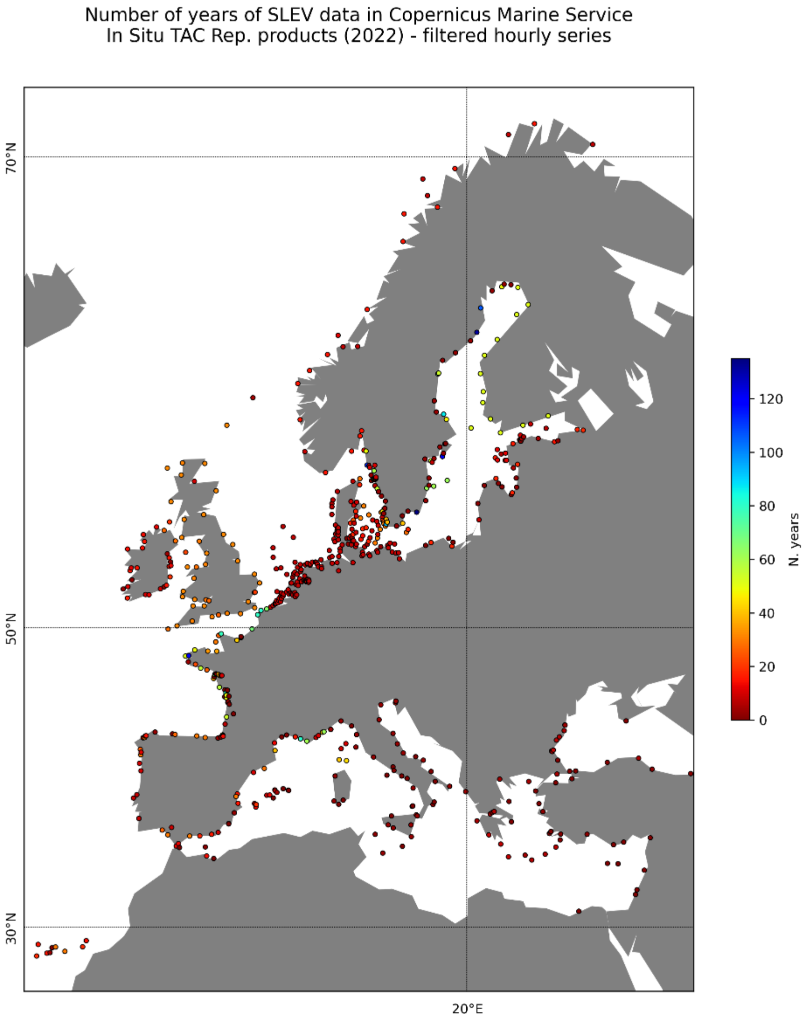


Figure 10: Stations with hourly reprocessed values in the sea level MY product, European region (colour scale number of years).

IV.3 Overview of data quality of the Sea Level REP product

An overview of the quality of the data in the product after the validation process is provided, and compared to the status of original data available in In Situ TAC NRT history folder.

From a total number of 1047 datafiles from 699 platforms in In Situ TAC NRT, 77 files (around 7%) and 19 platforms have been discarded for the November 2023 release of the product. These include those time series with length shorter than one year, due to too many incorrect data, too many gaps, or less than one year of data available until the end of 2022, the final date of the product.

MY product, original sampling dataset:

A comparison is made of the final percentage of observations flagged as “good” for the global MY sea level product, as compared to the NRT product. While 98.2% of data have QC=1 in the In Situ TAC NRT, the percentage increases to 98.1% in the new product, In Situ TAC MY, as a result of the delayed mode validation, with an overall decrease of flagged data of 0.3%. This results from the combination of different factors, including originally data flagged as dubious (QC=2) for the whole timeseries in some stations, which were finally flagged as good after the delayed mode validation process in the REP product. It is important to note that if the NRT data have been previously flagged by delayed mode processes of data providers, these flags are usually respected and maintained. Experience reveals that these are typically confirmed as good data during the delayed-mode validation in this product.

MY product hourly values dataset:

The quality of this dataset is analysed in terms of the annual percentage of valid data (not null values) or Completeness Index: $C.I.$ Table 6 provides an overview of this metric over the same periods and regions as those in Table 5, including average Completeness Index $\overline{C.I}$ and mean number of years with $C.I$ > 70%, considering all stations in a region and the corresponding period.

Hourly sea levels dataset	**Completeness Index C.I.**					
	1880-1950		**1951-1985**		**1986-2022**	
	$\overline{\boldsymbol{C.I.}}$	$\overline{\boldsymbol{Ny > 70}}$	$\overline{\boldsymbol{C.I.}}$	$\overline{\boldsymbol{Ny > 70}}$	$\overline{\boldsymbol{C.I.}}$	$\overline{\boldsymbol{Ny > 70}}$
GL	**92.8**	**35.0**	**78.7**	**12.7**	**81.1**	**12.6**
AR					98.1	12.2
BO	97.0	50.4	96.8	19.5	91.3	13.7
BS					36.7	
NWS	94.8	51.0	91.3	21.8	87.1	10.1
MED	85.8	23.0	62.9	5.5	85.5	12.8
IBI	88.3	15.6	78.0	12.1	86.6	17.5

Table 6. Mean Completeness Index ($\overline{C.I.}$) and mean number of years with C.I.>70% ($\overline{Ny > 70}$) for all stations in different regions and three periods: 1880-1950, 1951-1985 and 1986-2022. GL: global (all stations); AR: Arctic region, BO: Baltic Sea; BS: Black Sea; NWS: North-West-Shelf; MED: Mediterranean and IBI: Iberia-Biscay-Ireland region.

For the global region (all stations in the product), the $\overline{C.I.}$ is remarkably high during the first period (92.8%), decreasing to 78.7% and 81.1% in the last two periods. This reveals that despite the very small number of platforms available in the earlier times, these were very well maintained and their quality was very good. $\overline{C.I.}$ is generally good (over 80%) for most regions and periods, except in the Mediterranean Sea during 1951-1985 (62.9%), and the Black Sea for the 1986-2022 period (only 36.7%). $\overline{Ny > 70}$ is largest in the Baltic Sea and the North-West-Shelf region, for the period 1880-1950 (> 45 years). In recent decades (1986-2022) this value is higher in the Iberia-Biscay-Ireland region.

V SYSTEM'S NOTICEABLE EVENTS, OUTAGES OR CHANGES

Nothing to mention

VI QUALITY CHANGES SINCE PREVIOUS VERSION

This is the second version of this product, extending the temporal coverage from December 2020 to December 2022.

Up to new 60 platforms or tide gauges have been incorporated, most of them integrated by the Mediterranean Sea In Situ TAC NRT Production Unit.

All timeseries have been reprocessed and inspected in this second version, due to an update on the spike/outliers detection test, and a new automatic datum/drift detection algorithm based on the comparison with nearby stations.

VII REFERENCES

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

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

Pérez Gómez, Begoña, De Alfonso Alonso-Muñoyerro, Marta, Huess, Vibeke, Rickards, Lesley (2010). Near Real Time Quality Control and validation of Sea Level in-situ data within MyOcean. CMEMS-INS-SEALELEVEL-RTQC. https://doi.org/10.13155/74307.

Pugh, D.T. (1987). Tides, Surges and Mean Sea level. Chichester: John Wiley and Sons. 472 pp.

Horsburgh, K. J., and C. Wilson (2007). Tide-surge interaction and its role in the distribution of surge residuals in the North Sea, Journal of Geophysical Research, 112.

Pérez González, I., Pérez Gómez, B., Sotillo, M.G., Álvarez Fanjul, E. (2017). Towards a new sea level forecast system in Puertos del Estado. Extended proceedings. 8th EuroGOOS Conference. 3–5 Oct 2017; Bergen, Norway.