Manals and Guides 83
Intergovernmental Oceanographic Commission

Quality Control of in situ Sea Level Observations

A review and progress towards automated quality control

Volume I

United Nations Educational, Scientific and Cultural Organization
Intergovernmental Oceanographic Commission

Manuals and Guides 83
Intergovernmental Oceanographic Commission

Quality Control of in situ Sea Level Observations

A review and progress towards automated quality control

Volume I

UNESCO 2020

IOC Manuals and Guides, 83 Vol.1
Paris, May 2020
English only

Editing team:
Begoña Pérez Gómez (Chair, Puertos del Estado, Spain)
Thorkild Aarup (IOC, UNESCO)
Elizabeth Bradshaw (NOC, United Kingdom)
Julia Illigner (GFZ, Germany)
Andrew Matthews (NOC, United Kingdom)
Bill Mitchell (BOM, Australia)
Lesley Rickards (NOC, United Kingdom)
Peter Stone (NOAA, United States of America)
Matthew Widlansky (UHSLC, United States of America)

Contributors and providers of other information:
Ruth Farre (SANHO, South Africa)
Thomas Hammarklint (SMA, Sweden)
Aram Kim (KHOA, Republic of Korea)
Philip MacAulay (DFO/CHS, Canada)
Fernando Manzano (Nologin, Spain)
Marco Picone (ISPRA, Italy)
Octavio Gómez Ramos (UNAM, Mexico)
Oda Roaldsdotter Ravndal (Kartverket, Norway)
Scott Stephens (NIWA, New Zealand)
Joanne Williams (NOC, United Kingdom)

Technical editing: Thorkild Aarup and Elena Iasyreva

Acknowledgements:
The editing team would like to thank Philip Woodworth, Guy Wöppelmann, Tilo Schöne and Laurent Testut for very helpful comments and suggestions.

For bibliographic purposes, this document should be cited as follows:

UNESCO/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)

Published by the Intergovernmental Oceanographic Commission (IOC)
of United Nations Educational, Scientific and Cultural Organization (UNESCO)
7, place de Fontenoy, 75352 Paris 07 SP

TABLE OF CONTENTS

1. INTRODUCTION

The Global Sea-Level Observing System (GLOSS, https://www.gloss-sealevel.org/) is an international programme conducted under the auspices of the Intergovernmental Oceanographic Commission (IOC) of UNESCO. GLOSS aims at the establishment of high quality global and regional sea level networks for application to climate, oceanographic and coastal sea level research. The programme became known as GLOSS as it provides data for deriving the 'Global Level of the Sea Surface'. A major component of GLOSS is the 'Global Core Network' (GCN) of approximately 300 sea level stations around the world for long-term climate change and oceanographic sea level monitoring (Figure 1). The Core Network is designed to provide an approximately evenly-distributed sampling of global coastal sea level variations. GLOSS can be considered a component of IOC's Global Ocean Observing System (GOOS), and particularly as a major contributor to its Climate and Coastal Modules.

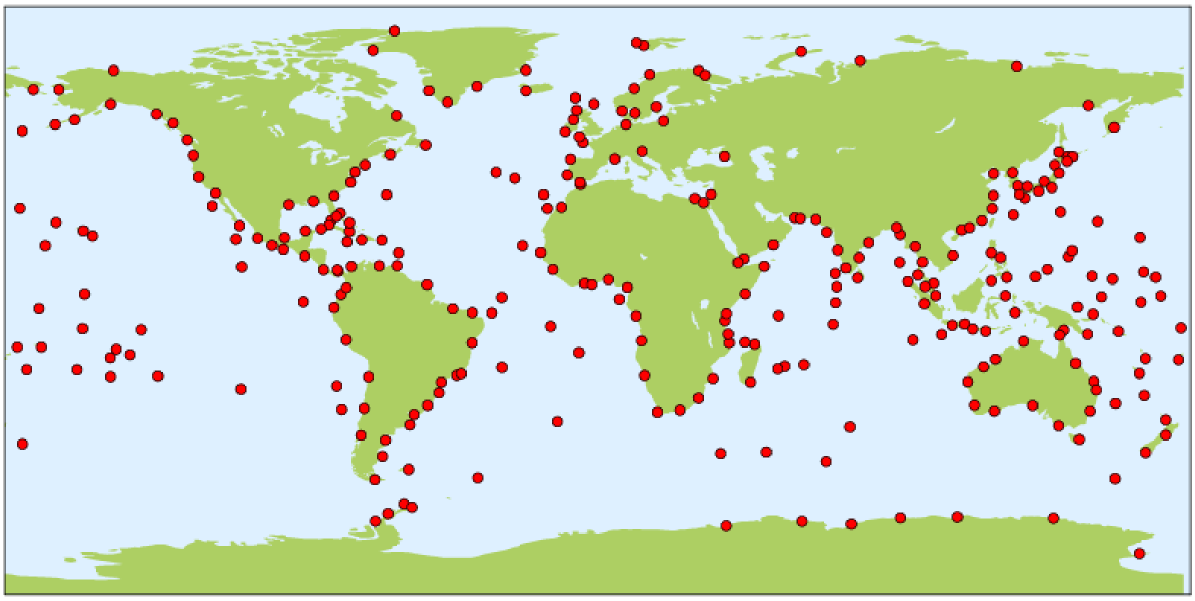


Figure 1.GLOSS Global Core Network (GCN2019).

In appreciation of the multiple uses of tide gauges, GLOSS has also sought to provide sea level data that meets the standards and requirements for tsunami warning and storm surge monitoring. Numerous GLOSS GCN stations have for many years contributed to the Pacific Tsunami Warning and Mitigation System (PTWS) and, following the 2004 Sumatra Earthquake, the IOC in consultation with GLOSS, has taken an active role in coordinating and implementing the sea level networks for the Indian Ocean Tsunami Warning and Mitigation System (IOTWMS), the Tsunami and other Coastal Hazards Warning System for the Caribbean and Adjacent Regions (CARIBE-EWS), and the Tsunami Early Warning and Mitigation System in the North-Eastern Atlantic, the Mediterranean and Connected Seas (NEAMTWS) (http://www.ioc-tsunami.org/).

Data from more than 70 countries are contributed to GLOSS, and in particular to the GLOSS Data Centres including the Permanent Service for Mean Sea-Level (PSMSL). However, quality control, although defined originally within the GLOSS programme, is accomplished at different levels and by different institutions and programmes at this moment. In addition, many sea level stations are not committed to GLOSS and the number of organizations dealing with tide gauge data (originators, facilitators and users) has increased. Thus, it is necessary to bring up to date the current good practice and distribute the information widely as a means of realizing a more standardized approach to quality control. Application of standardized sea level quality control, and agreed filtering techniques, will ensure that tide gauge data supplied to sea level data banks are consistent, and of a known accuracy. This will allow future researchers to better define confidence limits when applying these data.

The objective of this manual is to compile and update the standards and best practices on quality control of tide gauge data. Although related information has been included in the IOC Manuals on *Sea Level Measurement and Interpretation* (Volumes I (1985), II (1994), III (2002), IV (2006) and V (2016)), this is the first time that detailed information on these quality control procedures has been assembled into one document, addressing new issues like automation, for management of hundreds of long time series, or near-real time quality control procedures, for operational applications.

Before introducing the characteristics and objectives of the different specific quality control procedures, existing GLOSS data centres, related to tide gauge data flow and processing, are described below.

1.1. GLOSS DATA CENTRES

One of the main functions of GLOSS is to facilitate the smooth flow of sea-level data (and its associated metadata) from tide gauges to national and international data centres, and help provide scientists, operational users and others with efficient and seamless access to it. Stations with long time series from different countries throughout the world make up another GLOSS database termed the Long Term Trend (LTT) network (UNESCO/IOC, 1997; UNESCO/IOC, 2012) (Figure 2). Locations that have multiple years or decades of data have many applications in climate studies and other disciplines. It is essential that these datasets are quality controlled and processed in a consistent manner so that errors and biases are minimized.

Clause 1 of the IOC Oceanographic Data Exchange Policy states that “Member States shall provide timely, free and unrestricted access to all data, associated metadata and products generated under the auspices of IOC programmes” (see Annex I for the complete policy). This applies to data from sites in the GLOSS Core Network that Member States have committed to supporting. It should be stressed that GLOSS operates on a public service / public good basis.

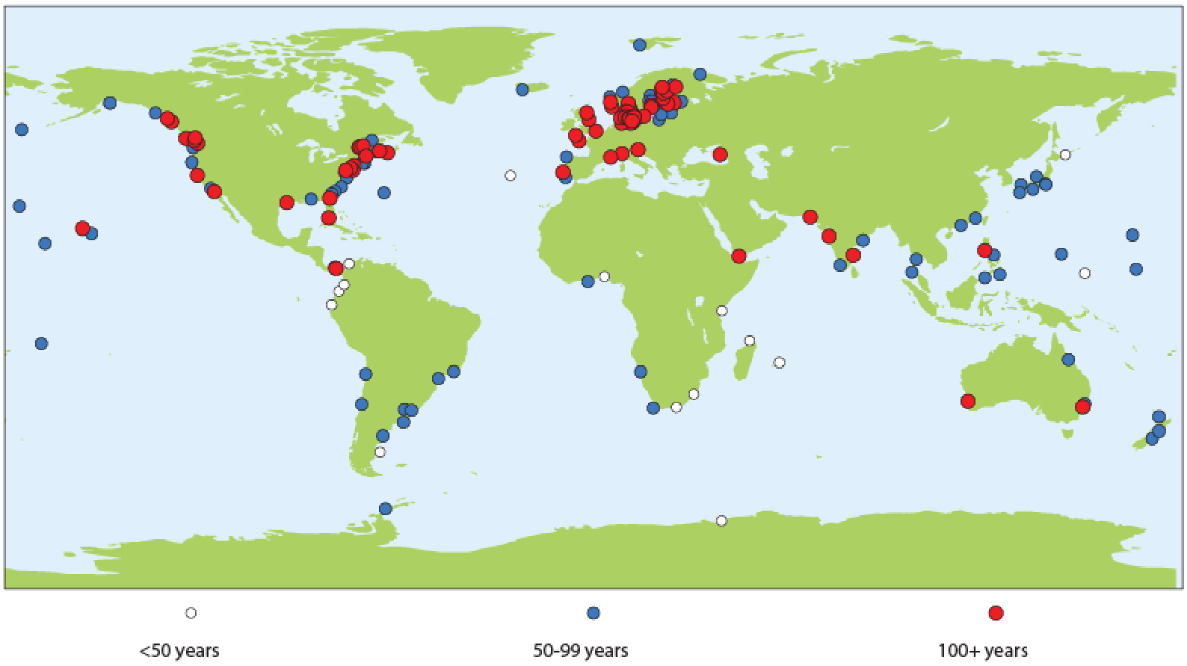


Figure 2. GLOSS Long Term Trend (LTT) network, with length of record available at each site, as defined in the GLOSS Implementation Plans (IOC, 1997; IOC, 2012).

The GLOSS sea level data centres described below provide complementary data streams while helping to shape the future of the global in situ sea level observing network. In general, GLOSS data centres are associated with scientists involved with sea level research, which helps to maximise the quality of GLOSS datasets. The centres also maintain close contact with data providers to ensure that information about changes in station location, instrumentation, or measurement datum can be communicated to users of the data.

Each of the data centres hold tide gauge data from many sites in addition to the GLOSS Core Network stations. Countries committed to GLOSS are required to provide data to these data centres:

(i) Permanent Service for Mean Sea-Level (PSMSL): Mean Sea-Level (MSL) data

Established in 1933, the PSMSL is responsible for the collection, publication, analysis and interpretation of mean sea level data from the global network of tide gauges, including the GLOSS Core Network (https://www.psmsl.org). It is based in Liverpool, United Kingdom, at the National Oceanography Centre (NOC). The PSMSL generally relies on Member Nations to provide the final version of the monthly time series with all quality control assessments applied and documented. Where possible, in order to construct time series of sea level measurements at each station, the monthly and annual means are reduced to a common datum using tide gauge datum history provided by the supplying authority. This provides the highest possible length and quality of sea level record at each location. In this way, the PSMSL archive comprises delayed mode monthly mean sea level values most suitable for studies of long-term sea level change: most studies of 20th-21th century global sea level rise are based on the PSMSL dataset.

(ii) GLOSS Delayed Mode Data Centre: delayed mode high frequency data and ancillary variables

The GLOSS Delayed Mode Data Centre (https://www.bodc.ac.uk/data/hosted_data_systems/sea_level/international/) is operated by the British Oceanographic Data Centre (BODC), UK, in collaboration with the PSMSL. It has the responsibility for assembling, quality controlling and distributing the “final” version of GLOSS sea level datasets, as well as all supporting metadata information (including benchmark details). The Delayed Mode Centre handles hourly (or sub-hourly) sea level measurements, together with ancillary variables (e.g. atmospheric pressure) where these are available, from the GLOSS sites. It generally relies on Member Nations to provide the final version of the hourly (or sub-hourly) time series with all quality control assessments applied and documented. On request from Member Nations, the Delayed Mode Data Centre will form monthly averages based on the final datasets received and provide them to the PSMSL.

(iii) GLOSS Fast Delivery Centre: fast higher frequency data

The University of Hawaii Sea Level Center (UHSLC), United States of America, operates the GLOSS Fast-Delivery Center (https://uhslc.soest.hawaii.edu/), which is responsible for assembling and distributing a version of GLOSS sea level data that has undergone preliminary quality control by Member Nations and includes supporting metadata information. “Fast-Delivery” implies making data available within 1–2 months. The UHSLC provides Fast-Delivery quality control services for Member States that do not have such capability. The Joint Archive for Sea Level (JASL) hosted by the UHSLC acquires hourly datasets, from GLOSS and non-GLOSS tide gauges from around the world that have received a final quality assessment from the data originators. JASL provides an independent check of the data, primarily to identify any remaining outliers, timing issues, or datum shifts. Any quality issues with the data are brought to the attention of the data originators for reconciliation. JASL then assembles a single hourly time series for each station or a series of sub-records if datum changes occur over time. The

JASL dataset therefore represents a “data product” as problematic data points are not simply flagged and left in the records as they are by the British Oceanographic Data Centre (BODC) for the GLOSS Delayed Mode Dataset but, rather, changes to the data are implemented by JASL analysts (e.g. level adjustments, timing shifts, outlier removal). These changes are documented in the metadata information.

(iv) IOC Sea Level Station Monitoring Facility (IOC SLSMF): real time data monitoring

The Flanders Marine Institute (VLIZ, Belgium) hosts the IOC Sea-Level Station Monitoring Facility that includes GLOSS core stations. VLIZ provides a web-based global sea level station monitoring service (http://www.ioc-sealevelmonitoring.org) for viewing sea level data received with variable latencies and samplings (approaching real time 1-min or less sampling data in some cases) from different network operators. Data are transmitted primarily via the Global Telecommunication System (GTS, global network for the transmission of meteorological data), a network protocol or satellite broadband. The service provides information about the operational status of GLOSS and non-GLOSS stations through quick inspection of the raw data stream. The sea level station monitoring system also runs a web-service for direct data access to raw data without quality control. The sea level station catalogue system developed and maintained at VLIZ links relevant station metadata information held at the different data centres.

(v) GNSS at Tide Gauge Data Centre (SONEL)

In 2001, the International GNSS (Global Navigation Satellite System) Service (IGS) set up a pilot Tide Gauge Benchmark Monitoring Project (TIGA), which set itself the task of processing and analysing continuous GNSS data from or near tide gauges around the world in a consistent global reference frame (see http://adsc.gfz-potsdam.de/tiga/ for more details). These data are essential for studies of sea level change, as it provides information about how the tide gauge is moving in three-dimensional space. The vertical movement of the gauge will register as local sea level change, but will mask global sea level changes. The main objective of TIGA is to learn more about the practical problems of using continuous GNSS in the coastal environment, to give advice to tide gauge station operators for setting-up and operating GNSS stations, and to collect and maintain a repository of GNSS observations at tide gauges and metadata (e.g. levelling information between the GNSS and tide gauge benchmarks) (Schöne et al., 2009). Since 2010, TIGA has been converted from a Pilot Project to a Working Group in recognition of its long-term importance. Several TIGA Analysis Centres reprocess GNSS data from long-term archives with the most recent software and methods to provide homogeneous and consistent geocentric coordinates time series and vertical velocities. In particular, TIGA works closely with the Système d'Observation du Niveau des Eaux Littorales (SONEL), France, to archive CGNSS (Continuous GNSS stations) data and distribute analysis products (http://www.sonel.org).

SONEL is the dedicated centre for GNSS data at or near tide gauge stations. SONEL is supported by the University of La Rochelle and the French CNRS/INSU institute. SONEL provides information about the status of GNSS stations at or nearby tide gauges through a web-based monitoring facility. It assembles, archives, and distributes GNSS observations and metadata that can be accessed through its web-based facility, as well as via an anonymous FTP server. SONEL provides easy access to GNSS products computed by TIGA analysis centres (and others that comply with the latest IGS reprocessing standards) at tide gauges co-located with a GNSS receiver, giving information about the geocentric height of stations, and their movement over time. This can help with interpreting the cause of a trend in relative sea level at the site during delayed mode quality control. (Wöppelmann et al., 2007, Wöppelmann and Marcos, 2016, Woodworth et al., 2017a).

(vi) Other tide gauge data centres

Nowadays, tide gauge data are also being distributed through other international data portals that usually deal with several geophysical variables: e.g. the Copernicus Marine Environment Monitoring Service (CMEMS) (http://marine.copernicus.eu/) or the European Marine Observation and Data Network (EMODnet) (http://www.emodnet.eu/). These gather data from original data providers or via agreements between international programmes. Another relevant initiative is the Global Extreme Sea Level Analysis (GESLA) dataset (Woodworth et al., 2017b: https://www.gesla.org), that compiled all tide gauge data from existing data portals and individual data providers to generate a new global dataset of tide gauge data ready to be ingested in scientific studies. At national and regional levels, tide gauge data are distributed and exchanged routinely for different practical applications, including tsunami and storm surge warning systems.

1.2. WHY DO WE NEED QUALITY CONTROL?

Data quality control essentially has the following objectives:

> "To ensure the data consistency within a single data set and within a collection of data sets, and to ensure that the quality and errors of the data are apparent to the user who has sufficient information to assess its suitability for a task." (UNESCO/IOC, 1993)

If done well, quality control brings about a number of key advantages:

- Maintaining Common Standards

All sea level data should be quality controlled to a minimum level. While archiving data just because they have been collected could prove useful for some applications or studies, the data should be qualified by additional information concerning methods of measurement and subsequent data processing. Definition of common standards, and distribution of guidelines on quality control procedures and required additional documentation, should be available to maintain the quality and long-term value of the data that are accepted, and to keep common standards to a higher level.

- Acquiring Consistency

Where data are held by more than one organization or data centre, the data should be as consistent as possible. It is possible that differences could arise between a real time and delayed mode version of the data, or where higher frequency data are stored at one location and hourly values at another. Any differences should be documented thus enabling external users to use the data with confidence. Searches for datasets are more successful as users are able to identify the specific data they require, even if the origins of the data are different on a national or even international level.

- Ensuring Reliability

Data centres, like other organizations, build reputations based on the quality of the services they provide. To be useful to the research community and to others, their data must be reliable, and this can be better achieved if the data have been quality controlled to a 'universal' standard. Many national and international programmes or projects carry out multidisciplinary investigations that require certain universal information on the marine environment. Many large-scale projects are also carried out under commercial contracts such as those involved with the hydrocarbon or fishing industries. Significant decisions are made, and theories formed, on the assumption that data are reliable and compatible; even when they come from many different sources.

In addition, quality control will contribute to alignment of tide gauge data with 'FAIR' Guiding Principles for scientific data management and stewardship (Wilkinson et al., 2016). These aim to ensure that data are ‘Findable, Accessible, Interoperable, and Re-useable’. In this regard, high quality and accurate metadata play an important role.

The quality control procedures include checking for unexpected anomalies in the time series, or in the derived tidal parameters, and in the filtering of the raw data to provide products such as hourly data, surge or non-tidal residual data, extreme sea levels and monthly mean values. These data and products are submitted to archiving centres such as the ones described in Section 1.1, for tidal applications, verifying satellite altimetry measurements, assimilation into numerical models, or climate-change related studies of sea level variations.

For such applications, the documentation of datum information (e.g. relationship of the recorded sea levels to the level of benchmarks on land) and history of the observing site are essential. Diagrams, maps, photographs and other metadata are also beneficial. Yet, there is at present little standardisation of methods for consolidating and archiving such information.

Today, quality control is also related to issues such as the availability of data in “real” time (real time definition is an issue, usually “near-real” time can be best applied to tide gauge data). If data are inspected every day or, in advanced systems, if data can be flagged for errors by automatic software with even sub-daily latency, then some faults can be rapidly attended to and fixed. Most importantly, such data can be automatically displayed or integrated in operational applications with higher confidence (Pérez Gómez et al., 2013). The quality control that can be automatically applied in near-real time to a window of recent data (usually a few days), is a preliminary assessment of the quality valid for operational applications. A second level of quality control, more complete and applied annually or to the whole historical time series, must be performed in delayed mode, to detect less obvious problems that can be checked with access to metadata before generation of the final products for that particular station. The latter, even when performed with the same automatic algorithms, will require some human intervention. Between these two levels of quality control, it is strongly recommended that the data are inspected every week or month, as the data come in, by a dedicated technician in close contact with the technicians in charge of the installation and maintenance of the station. This will allow them to identify and document faults more quickly. Finally, in some cases, as for tsunami warning, it is of utmost importance that the data are made available rapidly, and quality control/interpretation delegated to the expertise and knowledge of operators.

1.3. HISTORICAL REVIEW OF QUALITY CONTROL PROCEDURES

As mentioned in the introduction, this manual draws on existing publications and is adapted from frameworks developed by multiple programmes (World Ocean Circulation Experiment [WOCE] Sea Level Data Assembly Centres, the IOC’s International Oceanographic Data and Information Exchange [IODE], JCOMM Data Management Programme Area [DMPA] and the International Council for the Exploration of the Sea’s Data and Information Group Management [ICES DIG]), projects (e.g. European projects such as ESEAS-RI [European Sea Level Service, Research Infrastructure], SeaDataNet, MyOcean, AtlantOS) and regional and national initiatives over the past decades.

Manual/report/publication	Link
IOC Manuals and Guides No.14, Volumes I,II,III,IV,V (IOC, 1985, 1994, 2002, 2006, 2016)	https://unesdoc.unesco.org/ark:/48223/pf0000065061 https://unesdoc.unesco.org/ark:/48223/pf0000149528 https://unesdoc.unesco.org/ark:/48223/pf0000125129 https://unesdoc.unesco.org/ark:/48223/pf0000147773 https://unesdoc.unesco.org/ark:/48223/pf0000246981
European Sea Level Monitoring: Implementation of ESEAS Quality Control (García et al., 2007)	https://link.springer.com/chapter/10.1007/978-3-540-49350-1_11
WOCE Experience: Developments in Sea Level Data Management and Exchange. (Rickards L. J. and Kilonsky B. J. 1998)	https://www.bodc.ac.uk/projects/data_management/international/woce/documents/odspaper.pdf
SEADATANET: Data Quality Control Procedures (2010)	https://www.seadatanet.org/content/download/596/file/SeaDataNet_QC_procedures_V2_%28May_2010%29.pdf
EuroGOOS Recommendations for *in-situ* data Real Time Quality Control (Pouliquen et al., 2011)	http://eurogoos.eu/download/reference_documents_/rtqc.pdf
U.S. Integrated Ocean Observing System (IOOS), QARTOD, 2016. Manual for Real-Time Quality Control of Water Level Data: A Guide to Quality Control and Quality Assurance of Water Level Observations (IOOS, QARTOD, 2016)	https://cdn.ioos.noaa.gov/media/2017/12/qartod_water_level_manual.pdf
EU AtlantOS D7.2 Recommendations for an automatic RT or NRTQC for selected EOVs (T&S, Current, Oxygen, CHla, Nitrate, Carbon, Sea level) (G. Reverdin et al., 2017)	https://www.atlantos-h2020.eu/download/7.2-QC-Report.pdf
IOC's International Oceanographic Data and Information Exchange (IODE), Manual of quality control procedures for validation of oceanographic data (UNESCO, 1993)	https://unesdoc.unesco.org/ark:/48223/pf0000138825
International Council for the Exploration of the Sea's Data Guidelines for Water Level Data (rev. 2006)	http://ices.dk/sites/pub/Publication%20Reports/Data%20Guidelines/Data_Guidelines_TWL_v7_revised_2006.pdf

Table 1. Relevant publications containing quality control recommendations for oceanographic or sea level data.

As new data needs and applications arose, the quality control standards and procedures were upgraded. Table 1 contains a list of previous relevant publications and programmes considered here and available online. The objective of all these documents and efforts was to derive a set of recommended standards for quality control that will result in datasets acquired and processed to agree upon standards, which should be updated at regular intervals. Tide gauge data that meet such standards obtain the GLOSS quality endorsement.

Three basic data sources should ideally be available from every tide gauge station (Figure 3): raw observations, instrument calibration data, and levelling information. Sometimes ancillary data (e.g. atmospheric pressure) are also available and highly desirable. Collectively, these data are essential for the successful quality control of sea level data and production of derived products.

Most present day tide gauges have digital sensors, allowing sea level data to be collected as electronic records. Older stations produced graphical records, which need to be digitised subsequently. Raw data are usually produced at a particular sampling interval, generally more frequent than 1 hour (e.g. 5 minutes, 6 minutes, 10 minutes), averaging over a time period to cancel out the impact of waves on individual spot readings. Today, as an answer to tsunami warning requirements and the need for understanding high-frequency sea level oscillations and their contribution to extreme sea levels, raw data are commonly obtained every minute or even more frequently (1 sample/second or higher) in order to record and transmit tsunami wave-form data from all seismic and non-seismic sources. This change of sampling in raw data - the starting point of the data processing - has had an impact on the first level of quality control, as will be described in the next section.

From these raw data, hourly values (still valuable and sufficient for many applications) are obtained by means of a suitable filter, or by manual smoothing when digitizing graphical data, from which harmonic constants (that define the tide at the station), non-tidal residuals (observations minus tide predictions), and mean sea levels can be computed. Extreme values (i.e., observed highs, lows, and ranges between such) can be derived from hourly values or, if possible, from the higher-frequency (sub-hourly, varying depending on the station) quality controlled data. Again, the original sampling of raw data may have an impact on these last products. The data flow scheme of tide gauge data processing is represented in Figure 3.

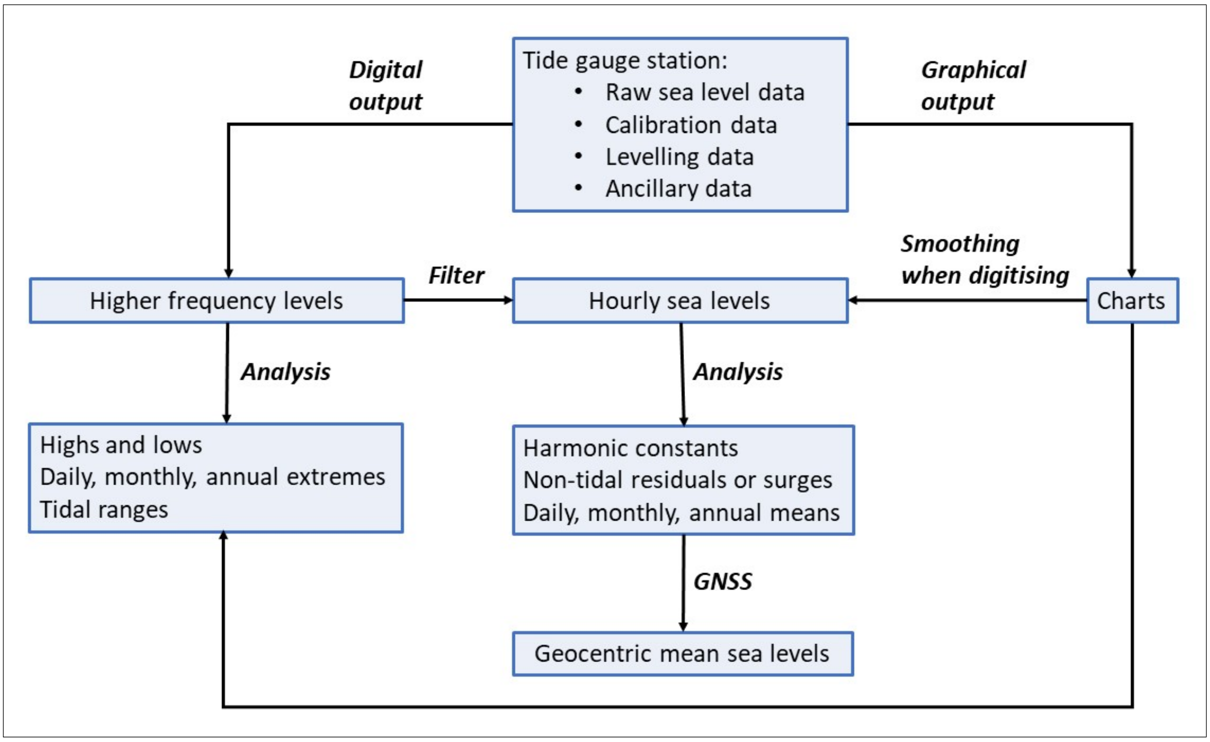


Figure 3. Scheme of relevant data available at a tide gauge station and the data processing normally applied to sea level data, for computation of the main derived sea level products. Most present day sensors will provide digital output. Old chart records are still available at many places from the end of the XIXth century.

1.4. TYPICAL ERRORS IN A TIDE GAUGE TIME SERIES

The errors that can arise in the sea level data and related parameters could be random errors (for example, transposition of numbers in manual recording or recording of observations in the wrong column), or could arise from electronic noise in measurements (spikes), problems in the communications (distortion or gaps), sensor calibrations (datum changes), floating objects below the sensor (in the case of radar sensors), clock malfunctions, drifts or stabilization (constant value). Some of these errors can today be detected and flagged by automatic algorithms, which can avoid clearly wrong data contaminating operational applications.

There can also be systematic errors or inhomogeneities in the sea level data, which arise primarily when there is a change in observational practice, or a change of instrumentation. Changes of sensor, location or both can result in sharp discontinuities in the sea level data, and have different effects on sea level processes on a variety of timescales. Changes in the environment surrounding the station, such as harbour construction, land movement, etc. can produce trends in the data or changes in the tide parameters. These types of errors require a more detailed analysis by a sea level expert. Access to metadata at this stage becomes crucial for successful quality control of the sea level data.

Based on these common types of problems, and on the general data flow and steps during data processing, recommendations on specific quality control algorithms and data processing methods, part of the quality control process as well, will be further described in Section 2.

2. RECOMMENDATIONS ON QUALITY CONTROL AND DATA PROCESSING ALGORITHMS FOR TIDE GAUGE DATA

The initial process of quality control consists of (i) performing various checks on the original series whilst maintaining the original data with flags that qualify them and (ii) creating a new modified time series that corrects or modifies some of the errors detected (e.g. interpolation). Although data may be modified as a result of a quality control process, the original series must always be preserved. For both series, the data should be flagged according to predefined quality control codes. Although specific values of quality flags may be used in the original raw data for identifying particular types of error, and this may be different depending on the organization, this is acceptable provided that they are defined.

Quality control of sea level data may be undertaken using automatic tests, by visual inspection and with more sophisticated analysis, combined at the different levels of quality control and processing as will be described below. For this reason this section will start with a summary of the different data streams and a brief outline of their associated quality control, followed by the description of specific tests and algorithms recommended for on-site checks, near-real time and delayed mode quality control and processing. Finally, some additional quality control techniques with particular emphasis on historical data, and recommended quality control flags, will be presented at the end of the section.

2.1 DATA STREAMS

Real time: For real time data provided as part of the tsunami monitoring system, with latencies under 1 minute, very little quality control is required or even possible. It is of prime importance that the data are provided without delay to the national and regional tsunami warning systems and, if possible, to the IOC Sea Level Station Monitoring Facility (http: www.ioc-sealevelmonitoring.org). Care must be taken to ensure that any quality control carried out on real time data does not remove tsunami signals by rejecting out of range data. A few simple checks may be carried out in real time: for example, detecting when the tide gauge stops reporting data, so that it can be fixed as soon as possible. At the tsunami warning centres, data should be checked by experienced personnel prior to entering any tsunami alert process. Further checks can be done in near-real time and delayed mode (described below) prior to archiving the data.

Near-real time (*Level 1 Quality Control - L1*): Data are considered to arrive in near-real time for latencies normally between 15 minutes and several weeks, and this is generally the requirement for latency in storm surge forecasting, operational oceanography, and in altimetry data validation. This longer latency allows the implementation of some level of automatic quality control prior to archiving and use of the data. L1 quality control consists of the detection of invalid characters, wrong assignment of date and hour, spikes, outliers, interpolation of short gaps, stability of the series and, depending on the application, even filtering to hourly values and computation of non-tidal residuals.

Delayed mode (*Level 2 Quality Control - L2)*: This is applied to long time series (one or more years of data), which require a more complete checking and analysing procedure, including computation of all derived sea level products such as harmonic constants, extremes (high and low sea levels, monthly and annual extremes), daily, monthly and annual mean sea levels and tidal ranges. At this stage, comparison with neighbouring tide gauges (“buddy checking”), altimetry data or ocean models can be useful, especially for allowing detection of drifts and datum changes and thus for studies of the long-term evolution of mean sea level. The knowledge of the operational history (sensor replacements, maintenance incidents) at the tide gauge station is essential in this case.

It can be seen that the generation of products (data processing) is part of the whole process of quality control because some errors are easier to identify in these products, allowing us to go back to the original time series and correct the problem in an iterative process. The two levels of quality control L1 and L2 were first defined within the ESEAS-RI European project (García et al., 2007). Further details on how L1 and L2 quality control procedures are actually implemented today, and the most common and useful algorithms, are described in Sections 2.3 and 2.4 below.

2.2 ON-SITE CHECKS, CALIBRATION AND QUALITY CONTROL

Before entering the sections on remote quality control of data, a brief description and examples of possible on-site checks should be taken into account. A more detailed description of on-site best practice (installation, sensor calibrations) and sensor types (float, pressure, radar) are outside the scope of this document and can be found in the existing GLOSS manuals on *sea level measurement and interpretation* (See Table 1).

One common way of estimating the quality of sea level measurements on-site, is to use the Van de Casteele test during maintenance visits. The procedure consists of taking manual readings (generally with a water level dipper whose accuracy is well known) simultaneous to the tide gauge readings during a tidal cycle (UNESCO/IOC, 1985). The two resulting time series will then be used to produce a diagram by plotting the sea level measured at the tide gauge (Y axis) against the difference between the two time series (X axis). Assuming that the manual readings are accurate and that the tide gauge is working properly, the diagram should be a straight line centred on the same X value (i.e. the difference between the two time series should be constant). Otherwise, the shape of the diagram will allow the detection and identification of different types of instrumental faults affecting the data quality (Martín Míguez, B. et al., 2008).

In an ideal case, a second previously calibrated radar sensor can also be temporarily installed during the maintenance visit and the Van de Casteele and other statistical tests (correlations, bias, root mean square errors) applied to the simultaneous time series. This approach is followed for example by Puertos del Estado (PdE) in Spain, for the Spanish REDMAR network.

Modern radar gauges output raw higher frequency data (1 sample/sec or higher). In these situations, on-site algorithms may provide another level of quality control before computing sea level data at several minutes time sampling. For example, NOAA sensors use an on-site algorithm that rejects individual samples, which fall outside the 3-sigma band, and then average the remaining samples to construct the 6 minute observation. In Australia, at the instrumentation level, all 1-second sea levels are averaged over a minute with standard deviations and outliers outside the 2-sigma band recorded and discarded, and then the mean re-evaluated. A similar approach is followed by the radar sensors used in the REDMAR network, in Spain, that also discards outliers outside the x-sigma (configurable) band in the 2Hz raw data before computing the 1-min sampling data.

2.3 BASIC QUALITY CONTROL TESTS AND DATA PROCESSING ALGORITHMS

The first step of quality control, applied in near-real time or in delayed mode, is the detection of most obviously incorrect data in the raw data stream by means of a combination of the following basic quality control tests:

- Invalid characters detection and syntax test: e.g. checks that the data file conforms to an approved format and that the file is the expected size to ensure the full message has been received.

- Checking for and flagging of out-of-range values (based on predefined upper and lower limits dependent on the station and included in the metadata for each station).
- Spline-fit algorithm for detection of spikes: fitting a spline to the data and discarding the values exceeding n-sigma from the fit. The fit is applied to a moving window, usually a few hours long.
- Stability/flatline test: adding a quality flag value when there is no change in the magnitude of sea level after a number of time steps. The number of data values or time steps to begin to flag depends obviously on the time interval, the sea level units (mm or cm) and the natural sea level variability. The number of similar consecutive data points allowed depends on the location of the gauge, the point within the tidal cycle and the sampling interval of the data.
- Date and time check ensuring the timing channel grows chronologically, preferably at a regular time interval, and check for gaps.
- Comparison of observations and tide predictions and some quality control applied to non-tidal residuals (e.g. timing errors appear as tides in non-tidal residuals).
- Rate of change test, that ensures a time series does not exceed a rate of change above a threshold value assigned for each site.
- Attenuated signal: checking for a series where the signal is not a flat line, but diminishes over a number of cycles. This can be if a well orifice is blocked, or if a compressor fails when using a bubbler gauge.
- Multivariate test: to compare the primary observation with a secondary parameter. This could include comparing the sea level data with atmospheric pressure or wind measurements.
- Neighbour test or "buddy checking", where data from one station would be compared with either a second sensor at the same location, or possibly with data from a nearby station to see if events propagate along a coastline.

The rate of change and position tests are standard tests often applied to other geophysical data. The latter may not be so interesting for fixed stations like tide gauges (it is critical for a buoy, however, because it allows drift detection, and would require the coordinates to be included in the record; see section on metadata). The last four tests listed above (rate of change, attenuated signal, multivariate test and neighbour test) are included and recommended in the QARTOD manual of NOAA (US, IOOS, 2016). The rate of change may also be difficult to apply to high-frequency sea level data (e.g. 1 minute sampling) due to the large natural variability of sea level at these time scales. In Australia, the National Tidal Unit NRT QC is based on the comparison of tide predictions and observations (residuals inspection), combined with some elementary range checking. For 1-min data, 1-min predictions are computed for each station one year ahead. In Spain, a combination of the spline-fit algorithm and non-tidal residuals inspection is done by applying the spline-fit iteratively to sea levels and non-tidal residuals.

The next steps of data processing and computation of products will allow a further analysis and detailed quality control for detection of less obvious problems.

Raw data are normally registered at time intervals between 1 minute and 1 hour. Before 2004, waves, seiches or tsunamis were not the main priority for GLOSS stations, as the focus was then on tides and mean sea levels. For this reason, the most common sampling for stations before then was 5, 6, 10 or 15 minutes. In UNESCO/IOC (2006) a recommendation was made to install new sensors with lower sampling (at 1 min or less) in those regions where higher frequency oscillations such as infragravity waves, meteotsunamis, seiches or tsunamis are to be detected. This new multi-hazard approach is reflected as well in IOC (2016), where the installation of two sensors, a radar gauge with 1 or 3 min average values (primary sensor), and a differential pressure sensor with 1 min or less sampling, is recommended. The pressure gauge would be the tsunami sensor and would enable short gaps in the primary sensor to be filled. It is important to stress that, up to now, according to GLOSS recommendations on quality control (e.g. chapter 8 of IOC (2016)) there is no need to undertake a complete QC of all 1-min data by GLOSS data centres. Advice is given, and reflected in this manual, to average these higher frequency data to those samplings commonly used by each institution, and to apply then the filters available to obtain hourly values. However, original 1-min data should be safely archived to be revisited and analysed for particular high-frequency events, or when new QC methods arise.

Nevertheless, it will always be necessary to deal with the transition from high-frequency sea level data to obtain filtered hourly values before proceeding with the rest of the sea level processing. Based on this, the next steps of tide gauge data processing (mainly described in detail in Pugh, 1987) are:

- Interpolation and resampling (average) of original flagged data. This generates a “clean” time series with flagged errors interpolated and homogeneous data sampling (< 1 h, usually 5, 6, 10 or 15 minutes), and a time series that can be entered into the next step of the process (filtering to hourly values).
- Filtering to hourly values: the filtering process will eliminate higher frequencies depending on the frequency cut-off. Pugh (1987) describes useful filters that can be applied to the sea level data at intervals of 5, 10 or 15 minutes to obtain the hourly heights whilst preserving the tidal phenomena. In Godin (1972) there is an extensive discussion on tidal filters.
- Harmonic analysis: computation of the tidal characteristics (amplitude and phase of the main harmonic constituents) at the location. This is usually performed on an annual basis.
- Computation of the tide and non-tidal (or surge) residuals: tide prediction based on the previously obtained harmonic constants for the data period, and computation of the non-tidal signal by subtracting the tide from the observations.
- Computation of tidal ranges, extreme sea levels, daily and monthly mean sea levels.

Each of these elements will be further explained below for near-real time and delayed mode.

2.3.1 Near-real time quality control procedures (L1 Quality Control)

A first level of quality control and processing can be easily applied automatically to tide gauge data for operational applications. Most institutions apply the same type of algorithms at this stage to flag most common random errors, implemented in different ways according to their internal needs, data flow and infrastructure (e.g. NOAA (US), BODC (NOC, UK) or the Australian National Tidal Unit). Puertos del Estado (PdE, Ports of Spain) developed software that implements this procedure automatically and that was later adopted by the MyOcean project and CMEMS (Pouliquen et al., 2011). For this reason this software (named SELENE [SEa LEvel NEar-real time quality control and processing]) will be used here to illustrate the main procedures, steps and algorithms, with mention of other tools applied by other institutions that, equally valid, may be slightly different in some cases.

The general flow of the near-real time process as used today by SELENE software is shown in Figure 4. The code is divided for simplicity in the following different modules:

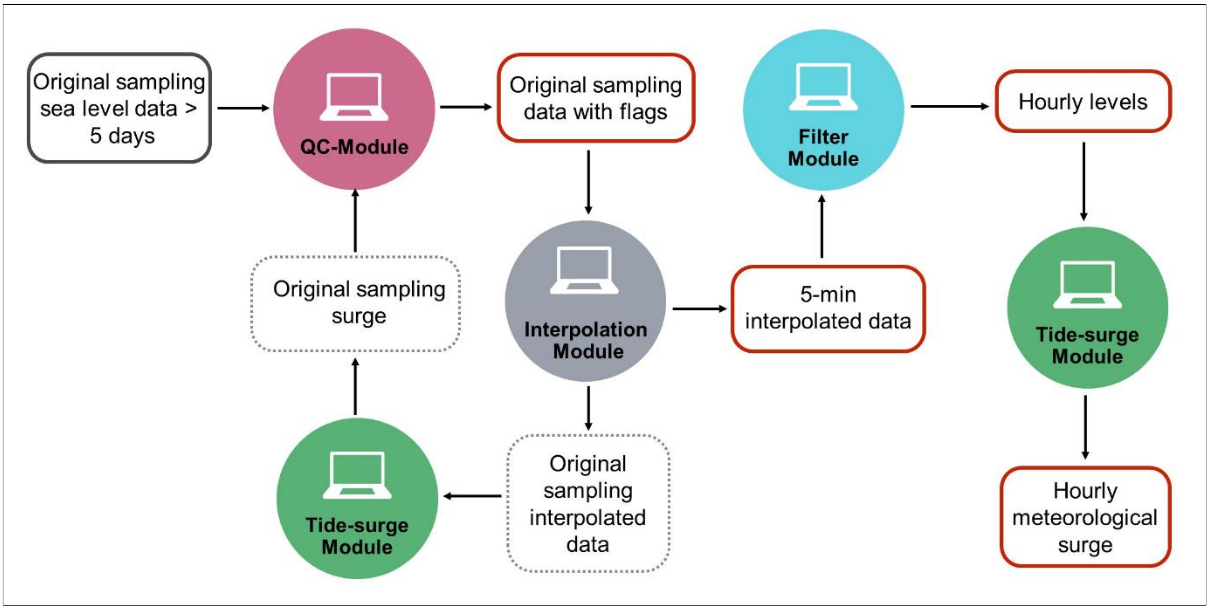
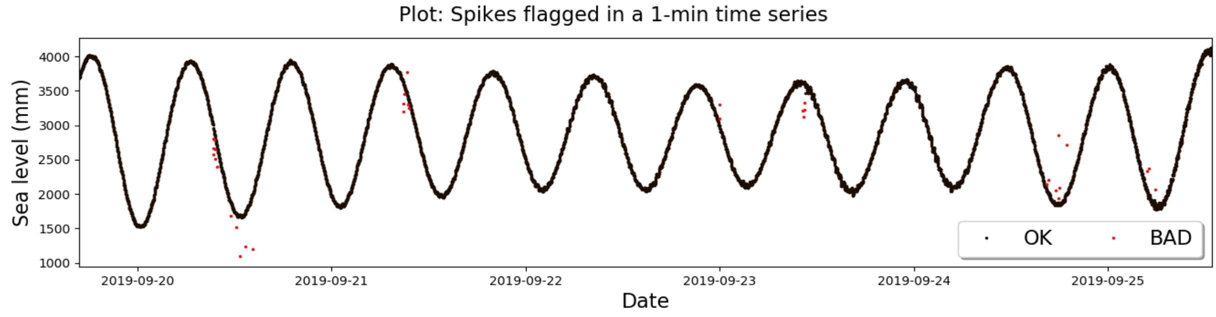


Figure 4. Scheme of the automatic software for quality control in near-real time now in place at Puertos del Estado (PdE) and used by SELENE software. Circles: different modules of the quality control and processing, red rectangles: consolidated products to be stored/distributed, dashed-line rectangles: intermediate products.

(i) QC module

The algorithm for the detection of spikes is the main component of the QC module. It is based on the fit of a spline to a moving window of around 12–16 hours. This cannot be applied in real time (latencies of 1 minute) because the procedure needs a long moving window to be able to detect spikes correctly, and not flag real phenomena such as sudden high frequency oscillations due to “seiches” or tsunamis. The degree of the spline (normally 2) and the size of the window can be selected depending on the characteristics of the tide, the data sampling, etc.

The algorithm flags as spikes the values that differ more than N standard deviations from the fit (normally N=3, although this can also be selected in the configuration file). Repeating the process for non-tidal residuals is crucial to detect less obvious spikes not detected in the first step; this is why the QC module is applied again when the residuals are obtained (Figure 4).



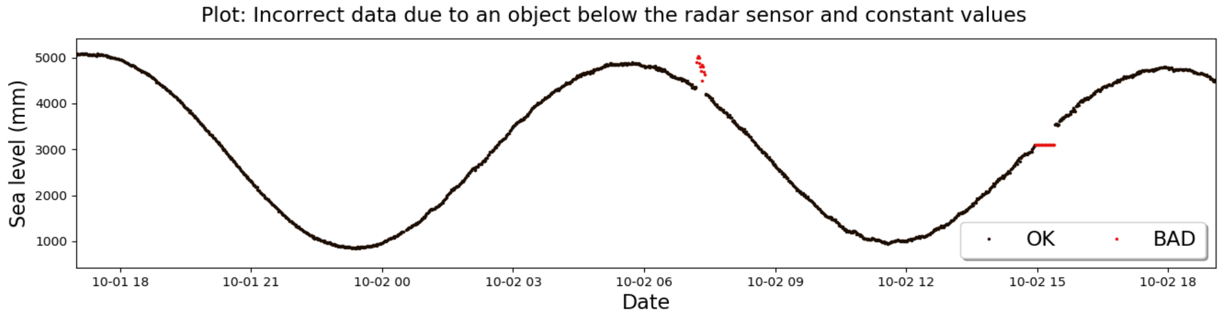


Figure 5. Examples of incorrect data (plotted in red). Top: output of the fit of the spline method to a tide gauge time series with spikes; bottom: example of the impact of an object below a radar sensor (first set of data flagged in red) and constant values (second set of data flagged in red). Sampling of data: 1-min.

Figure 5 (top) shows a typical example of the output of this algorithm for a time series with many spikes. Figure 5 (bottom) also shows examples of other types of errors that can be present and most of the times easily identified by this first step of the automatic quality control: an object located below a radar sensor that temporarily displays higher sea levels than expected. Depending on the duration of this problem, this will be correctly flagged or not, as will be shown in Section 4. Another type of error easy to flag and discard before forward data processing, is the presence of constant values, as can be seen in the same figure.

(ii) Interpolation module

Most of the raw data from a tide gauge arrives with a data sampling of several minutes, although for many applications in operational oceanography 1 hour is considered enough; besides, this data sampling is not always regular or, for example, 5 minute data supposed to arrive at 00, 05, 10… start arriving at 02, 07, 12. This is just an example of what can be found in the raw data.

The interpolation module includes the following functions:

- Checking and adjusting the time interval;
- Interpolation of wrong values previously flagged in the QC module;
- Filling the gaps with new records with the correct date assignment and special value for null-values;
- Interpolation of very short gaps (less than 10 – 25 minutes, ideally configurable depending on the tidal range).

This module has therefore as its main objective to homogenize the temporal sampling of the time series and, when the sampling is really small (1 minute or less), filtering the higher frequency oscillations (simple average) before generation of a 5-min (or 6-min, 10 min, etc., depending on the institution) time series. The output is a “clean” interpolated time series, ready to enter the filter and harmonic analysis programs, i.e. it will be the one used for the rest of the data processing.

The following steps are considered optional in previous quality control recommendations for near-real time applications (Pouliquen et al., 2011).

(iii) Filter module

Hourly values must be computed by means of a suitable filter, depending on the original data sampling. For example, for 5-min data, the following symmetrical filter can be applied:

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

where M=54, $X_f(t)$ is the hourly filtered value and $F_{0...m}$ the weights applied to the high frequency values. Details can be found in Pugh (1987).

This filter is for example used in PdE and is also one of the recommended filters found in the Copernicus Marine Environment Monitoring Service (CMEMS), EuroGOOS and ESEAS-RI Quality Control manuals (García et al., 2007, Pouliquen et al., 2011). The selection of the filter is open provided it is able to discard frequencies larger than 0.5 cycles/hour.

(iv) Tide-surge module

The last stage of the near-real time quality control may be the computation of the astronomical tide and the surge or non-tidal residual for the window of data. In SELENE software, this is performed by means of the Foreman method for tidal prediction (Foreman, 1977), and it requires the availability of the main harmonic constituents at each particular station, obtained in delayed mode from ideally 1 year of data. Access to these previous data is necessary in order to compute a reliable set of harmonic components.

Once the first non-tidal residuals are computed, the QC module is applied again to non-tidal residuals or surge data (Figure 4), in order to detect less obvious spikes. If detected, these newly-identified incorrect values are flagged again in the total sea level time series and the rest of the process is repeated to obtain the final products: interpolated series and hourly levels, surge and tide. Then the time series is ready to enter, for example, a storm surge forecasting system.

An example of the advantage of this second step of the automatic quality control, that requires additional metadata (tidal constants), is shown in Figures 6 and 7, for SELENE software. The availability of tide harmonic constants allows us to apply the second step of the software, as illustrated for a problem found in this time series, on 29 September 2019, when a sudden change of reference during the retreating tide led to wrong data values for several hours.

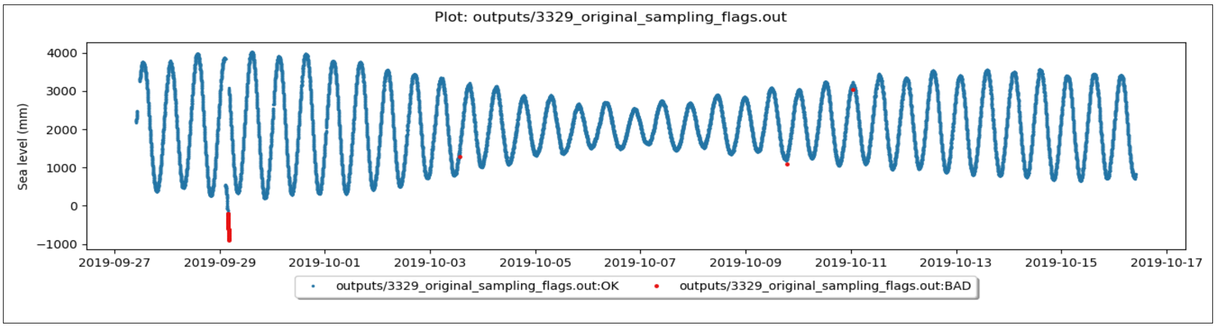


Figure 6. Output of SELENE software with only step 1 (no tidal constants available): good data in blue, flagged (wrong) values in red. Only a few data during a change of reference on 29 September are flagged with just this first step.

As can be seen in Figure 6, the application of this first step only, without tidal constants, would result in flagging just a subset of incorrect data. The software is not able to identify the whole period as wrong. However, in Figure 7, the outputs of the two steps of the quality control are displayed: in the top panel, the original data are flagged correctly: part of the data not flagged in the first step were flagged during the application of the spline-algorithm to the non-tidal residuals. Notice that the data are flagged and therefore discarded (gap too long for interpolation). Correction of the reference change (which could be possibly done manually), is not performed by the software. However, a gap in the final product (depending on the application, especially in near-real time) is less detrimental than the original wrong data values.

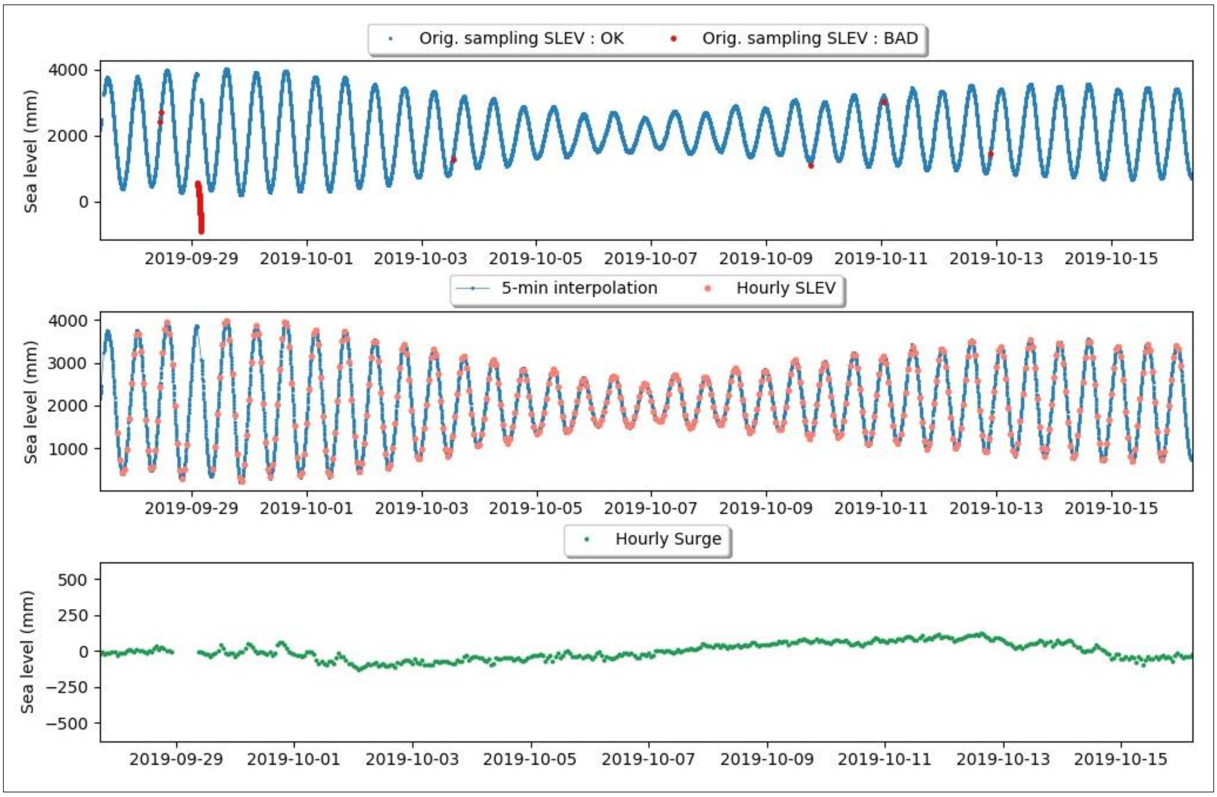


Figure 7. Output of automatic quality control with steps 1 and 2 applied (tidal constants available). Top: original raw data flagged, good data in blue, flagged (wrong) data in red; middle: by-products of the software: 5-min interpolated data (blue) and hourly filtered data (orange); bottom: by-product of the software: hourly non-tidal residual for the period (green). All wrong data have been correctly flagged in the top panel due to availability of tidal constants.

The remaining sea level products and data processing are usually obtained in delayed mode, as will be described below.

2.3.2 Delayed mode or 'scientific' quality control (L2 Quality Control)

As part of the "scientific" or delayed mode quality control (L2), more detailed processing of sea level data is performed, applied to longer time series (typically 1 year) that include not only the steps described for L1, but also computation of annual harmonic constants, non-tidal residuals, extremes and means. The results of this process are themselves useful products from the station, but also the examination of their quality is crucial for the detection of problems and malfunction in the tide gauge. The primary quality control of sea level is based on the visual inspection of both recorded data and non-tidal residuals, especially useful for detecting instrumental faults such as timing errors, datum shifts and spikes (Woodworth et al., 2015). Automated quality control software (as the QC module in Section 2.3) can also be used to aid and complement visual inspection of the data in delayed mode.

When comparing water level observations with the predicted tide levels, it is important to consider the quality of the tidal prediction. For example, the harmonic constants may be severely corrupted if the site is characterized by highly nonlinear tides, influence of rivers or estuaries and particularly complex basin configuration. To produce more accurate predicted tides, it is advisable to compute 'fresh' tidal constants from recent data (typically each year) and not simply rely upon historical values. Tidal analysis can be performed by means of different software packages: the already mentioned Foreman tidal analysis and prediction programs of the Institute of Ocean Sciences, Victoria, British Columbia (Foreman, 1977), used for example by SELENE software in PdE (Spain), descendants of the TASK Package, used by the PSMSL/National Oceanography Centre (UK) or TIRA tidal analysis programs (Murray, 1964) used by the Australian National Tidal Unit. More recent packages are the T_Tide/UTide tools (Pawlowicz et al., 2002, Codiga, 2020), available in Matlab and Python (https://pypi.org/project/UTide/), and adopted recently by the University of Hawaii Sea Level Center for delayed mode quality control. Whatever software is selected, tidal constants used in tide predictions should never be mixed between different packages.

An alternative approach to quality control that does not require comparison to tidal predictions involves cross-comparison among multiple water level observing sensors, e.g. records from a nearby site (also known as "buddy checking"). In case of a fault, data should be corrected or interpolated, if possible, otherwise the data must be maintained unchanged and the event noted as part of the metadata of the station. For this reason, it is recommended that more than one, ideally at least three, sensors are operated at the same site in order to allow direct comparison, and on occasion to fill gaps. This approach is followed, for example, by the Canadian Hydrographic Service (CHS).

Another example of this is the UHSLC use of "switch data" during delayed mode quality control. Every UHSLC station has at least one mechanical switch positioned in such a way as to be triggered (opened/closed) during each daily tidal cycle. The stations transmit times of the "switch" events. Since the switch elevations are carefully measured during maintenance visits, part of the quality control procedure is to analyse sensor water level measurements (e.g. from the radar instrument) at the times of opening and closing switches. If the sensor transmits a water level at the switch time that is different from the switch elevation, then this analysis reveals that either (1) there is a timing error with the switch or sensor, (2) the sensor is malfunctioning or moving (e.g. a pressure sensor drift), or (3) the switch is malfunctioning (bio-fouling of partially submerged mechanical devices is often a problem). The latter possibility is the primary source of weakness concerning this quality control option; however, cleaning of the switches on yearly or more frequent basis mostly mitigates such a problem.

Other quality control procedures can be applied in delayed mode, as described below:

(i) Review of harmonic constants variability

A common procedure is to compute the tidal harmonic constants for each year of observed data. Some of these constants may be particularly affected by meteorological conditions, and so will show important variations from one year to the next. This occurs for example for the long-period harmonic constituents such as Sa and Ssa. Sometimes the presence of problems in the data series appears as strange values of the normally stable harmonics (e.g. clock errors). In any case, an inspection of the variation through the years of the harmonic constants is interesting, both for detecting problems and also for providing information about station changes. For example, changes in the configuration of a harbour can affect the tide parameters.

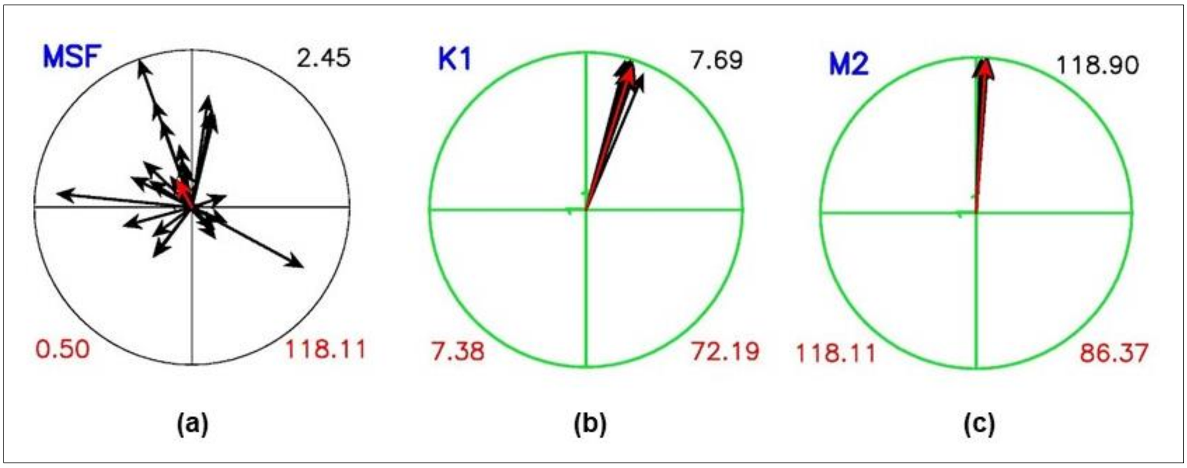


Figure 8.The vector representation is useful to observe the annual variations of the harmonic constants: examples shown: Figure 8.a for the constituents MSF (fortnightly tide, period: 14.77 days); Figure 8.b for K1 (lunisolar diurnal tide, period: 23.94 h); and Figure 8.c for M2 (principal lunar semidiurnal tide, period: 12.42 h), obtained from a particular tide gauge during 1993 to 2017. In black the arrows representing the amplitude and phase of each year and the maximum amplitude value (top-right); in red the mean vector and its amplitude (bottom-left) and phase (bottom-right) values. In this representation, green circles are used for those constituents showing little variability from year to year.

A useful representation of tidal harmonic changes with time is to calculate key harmonic constants on a yearly basis and plot their characteristics as vectors, as seen in Figure 8. In order to choose adequate harmonic constants for tide prediction, one can compute the vector mean and statistics of the annual values for several years (provided they are computed for nearly complete years and so the same number of constituents have been resolved) and select for prediction only the mean of those constituents which do not present a variability above a fixed and reasonable tolerance.

(ii) Inspection of non-tidal residuals

Inspection of non-tidal residuals is, as already mentioned, an essential tool for detecting errors during the quality control process. Most of the types of errors that a sea level time series can present are easily detected in the non-tidal residual plot (Pugh, 1987, Woodworth et al., 2015). An example of the presence of a clock malfunction (oscillations in the residuals) and reference changes can be observed in Figure 9.

Data spikes are typically obvious in the residual series as well, which is why some of the automatic algorithms for the detection of spikes are based on the inspection of the original and non-tidal residuals data.

(iii) Correction of clock malfunction

Identifying and correcting timing errors range in complexity from simple to involved, depending on the nature of the clock malfunction. Errors are often easy to correct if there is a constant time shift. The problem becomes more difficult to solve when there is a temporal drift in the lag between the observed and predicted tide.

Apart from the inspection of the residuals, where this error will cause an artificial periodic signal (Figure 9, left), a constant lag can be exactly determined by means of lag-correlation analysis between observed and predicted data (lag of maximum correlation), or by comparing the values of the phase of the M2 harmonic before and after the shift. If multiple water level measurements are available from different sensors with independent clocks, cross-comparisons are also useful.

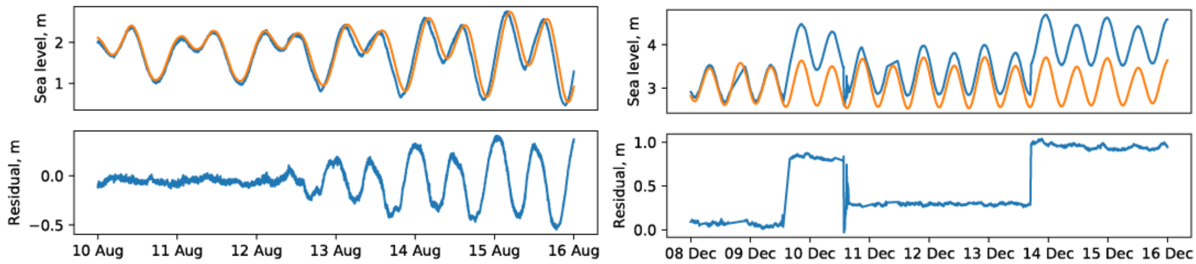


Figure 9. Left: Example of the trace of a clock malfunction in the non-tidal residuals (bottom plot); right: example of reference jumps in the non-tidal residuals (bottom plot). In both figures: top plots show the astronomical tide in orange and the observed total sea levels in blue. These types of errors are not corrected by existing automatic software packages.

Once a timing error is determined, the part of the series that is affected must be shifted accordingly to make the correction. If the lag is a multiple of the time interval, the shift is just a movement of data in time. Otherwise, an interpolation to the correct time has to be performed. These types of corrections are not typically implemented by automatic software packages. Instead, the analyst performing the quality control usually adjusts a set of software commands to correct for the particular timing error. Modern tide gauges including a GNSS receiver or internet connection for automatic time adjustment practically avoid this type of error.

(iv) Gap filling

Depending on the application, filling gaps in a series may or may not be reasonable. According to GLOSS manual II (UNESCO/IOC, 1994), in the section describing the TOGA Sea Level software, gaps less than 24 hours can be replaced by data from auxiliary gauges or by interpolation, if needed. This is also suggested as a possibility for correcting individual incorrect points (spikes) and glitches (1–6 consecutively obviously wrong points). In fact, the UHSLC interpolates gaps of less than 24 hours before computing daily and monthly means; this is done by computing the residual series, linearly interpolating by using the residual values at the extremes of the gap, and adding on the astronomical prediction to the interpolated values.

There was no specific recommendation for higher-frequency data (minutes) in the first GLOSS manuals, so this possibility of interpolation was applied basically to hourly values. However, during the first stage of quality control of the higher frequency data (1 min or less), very short gaps of a few minutes can have an impact on the filter to hourly values. For this reason, it would be useful to interpolate single point spikes and short gaps of a few minutes before entering the filter to hourly values. Interpolation of this kind should be undertaken with great caution, taking into account the natural variability of the time series, and the data values must be flagged accordingly.

(v) Detection of reference level changes

Reference level changes, which can appear in tide gauge water level measurements either abruptly or gradually (i.e., over hours or years), occur when the sensor moves vertically relative to the surface of the ocean. Such changes happen naturally (e.g., regional land motion caused by an earthquake) and artificially (e.g., sensor motion caused by an accident, improper maintenance, or localized land shifts). Many reference level changes are easily identifiable in hourly non-tidal residual plots, especially if the change is large and abrupt relative to the tidal cycle (e.g., Figure 9, right). Identifying smaller or more gradual reference level changes, perhaps due to a malfunctioning sensor or land motion, require careful comparison between long segments (i.e., months to years) of the tide gauge data with that from reliable neighbour stations, vertical site surveys potentially including GNSS measurements, and/or satellite altimetry measurements of sea surface height. Comparisons of monthly mean water level data are typically sufficient for identifying long-term changes of the order of a few centimetres.

Ultimately, the tolerance for what magnitude of reference level shift is unacceptable depends on the configuration and robustness of the tide gauge station. In general though, any change in reference level that is identified and well understood should be corrected and also documented in the station metadata.

(vi) Statistics

Basic statistics from historical data are computed or updated annually and some of these parameters are used for the quality control process. For example:

- Upper and lower limits of historical sea levels (for range check);
- Tidal and observed ranges (high to low tide);
- Maximum and minimum values, mean and standard deviation of hourly values, non-tidal residuals, ranges or mean sea levels;
- Tables of monthly and annual extremes;
- Density function for hourly values, tide predictions and residuals.

2.4 QUALITY CONTROL AND PROCESSING OF ENTIRE RECORDS

When working with historical data, even if the station is well documented, check sheets may not be available with which to perform a confident quality check on the reference level. Furthermore, system measurement problems, changes in the instrumentation or in the environment surrounding the station can generate a discontinuity, which may appear as a datum shift or a trend. In this case, some additional checks should be performed on the data. The normal procedure for this kind of higher level quality control is to work with several daily or monthly mean sea level series from nearby stations and then reconstruct the time series of the heights.

In addition to the more immediate computation of differences between levels of adjacent stations, which may clarify the reason for a problem, there are other algorithms (briefly described below) that can also help to detect this type of discontinuity or reference problem in historical data. All of them require the quality assessment of an expert before taking the final decision to correct the data.

(i) Correlations

Pearson's correlation coefficients can be computed both between data from different stations or sensors and between different parameters at the same station (wind, atmospheric pressure, etc), in any case a valuable tool for detecting problems and even for filling gaps. For the latter, the procedure would be the following: if the correlation of non-tidal residual series with a nearby station is above 0.7, a linear regression can be fitted between them and the regression fit be used to fill the gaps (if these are not at the beginning or the end of the series).

(ii) Standard Normal Homogeneity Test (SNHT)

Several tests have been described in the literature, which can be used to detect inhomogeneities in data series. Alexanderson (1986) developed the Standard Normal Homogeneity Test (SNHT), which is widely used in climatic time series studies. The SNHT gives the points where an inhomogeneity exists and provides information about the probable break magnitude. However, the inhomogeneity could be due to an error or to an anomalous, but real, behaviour of the variable. For this reason, the series are only corrected following comparison with other series in the same climatic region and supported by historical information about other instances in the tide gauge record.

(iii) EOF Analysis

The Empirical Orthogonal Functions (EOFs) analysis applied to a group of time series stations can be used not only to find spatially coherent signals or regional variability but also to detect possible errors in the time series. In fact, relevant differences in the variance of the first EOF may indicate errors in one or more time series. This technique is well documented in Marcos et al. (2005).

2.5 QUALITY CONTROL OF MONTHLY AND ANNUAL DATA

(i) Calculation of monthly and annual means

The calculation of monthly and annual averages from a sea level record can assist in identifying problems at a site, in addition to providing a valuable record of long-term sea level change. The PSMSL recommend calculating monthly means from so-called daily means computed by passing hourly means through a filter (such as the 39-hour Doodson, a 71-hour Demerliac filter or the filtered applied by UHSLC) that removes the tidal energy at diurnal and higher frequencies (UNESCO/IOC, 2002). A monthly mean can then be calculated as an arithmetic mean of the daily values. The number of missing days in each month should be reported, and the PSMSL recommends discarding the month if over 15 days are missing.

The PSMSL calculates annual means as a weighted mean of the monthly values, with each month weighted by the number of days present. If over one month is missing, the annual mean is not calculated.

Historically, monthly and annual means were sometimes reported using the Mean Tide Level (MTL), a value that was easier to measure and calculate before the era of automatic recording gauges. This is calculated by averaging the height of all the high waters and all the low waters in a month, giving mean high waters (MHW) and mean low waters (MLW) respectively. The MTL is the average of these two quantities. Great care must be taken when comparing MTL and MSL as they can differ by many centimetres, particularly in shallow water locations (Woodworth, 2017).

(ii) Quality control of monthly and annual means

A time series of monthly or annual data can highlight problems with the vertical datum control that are sometimes masked by tidal signals in higher frequency data. Potential problems include small datum jumps between sections of data separated by a gap when the gauge was not operational, or when there has been a change of technology to upgrade the station. Similar quality-control techniques as applied to higher frequency data can be used, such as using spike detection, or comparison with nearby sites. The analyst can use information from metadata to guide them: for example, particular attention should be paid to times when a station was relocated, or had new equipment fitted.

Comparison between the tide gauge data and measurements of sea level derived from satellite altimetry data provides an alternative approach to quality control. Satellite versus tide gauge comparison is especially necessary when there are not suitable nearby tide gauges for analyses (e.g., too far away or limited overlapping times). Typically, monthly (or, possibly daily) satellite data nearest the tide gauge station are used for the comparisons (Vinogradov and Ponte, 2011; Pérez Gomez et al., 2014; UNESCO/IOC, 2016). These data are provided as gridded products by altimetry data providers, from the combination of the different satellite altimetry missions since 1993 (e.g. http://marine.copernicus.eu/services-portfolio/access-to-products/). However, some care needs to be taken, as altimetry products become less accurate near land, and ocean processes between offshore altimetry data and the tide gauge position may be present, making these data not strictly comparable (comparisons at daily, monthly, or lower frequency mitigates some of these issues). Moreover, if long-term trends are to be compared, it should be noted that altimetry will not be affected by land movement that may be affecting tide gauge data. Another important drawback is the need to take into account the corrections applied to altimetry data, such as the atmospheric correction (wind and pressure), that will have to be added to the altimetry data before comparison with tide gauge monthly means (such atmospheric effects are removed by default from most altimetry products). Fortunately, this atmospheric correction (DAC: Dynamic Atmospheric Correction) is a product provided together with the gridded sea level anomalies by altimetry data providers (e.g. https://www.aviso.altimetry.fr/en/data/data-access.html). Both DAC and gridded sea levels must be daily or monthly averaged prior to the comparison with daily and monthly mean tide gauge measurements.

As described in UNESCO/IOC (2016), the Joint Archive for Sea Level at the University of Hawaii routinely compares daily means from tide gauges and gridded altimetry products (see Figures 8.a and 8.b in this IOC manual). The Permanent Service for Mean Sea Level has sometimes followed this approach for detection of severe datum problems. Pérez Gómez et al. (2014) used this technique for gross error detection in monthly mean sea levels after renewal and upgrade of the Spanish REDMAR network, allowing detection of differences between old and new technologies and even detection of hardware problems in some new radar antennas. Figure 10 shows an example of the impact of change of instrumentation on the monthly means for Ibiza tide gauge (Balearic Islands, Spain). The consistency between tide gauge and altimetry data improved when the new radar sensor was installed at the end of 2009, not surprisingly due to the uncertainties in the old pressure sensor affected by water density variations (details in Pérez Gómez et al., 2014). In general, high correlation between tide gauge- and satellite- measured sea level anomalies can be expected, especially on monthly time scales (Vinogradov and Ponte, 2011).

It must be emphasized that when it comes to changes in tide gauge zero, only substantial changes can be identified by the approach described above. For a more accurate evaluation of drifts, some authors suggest a direct comparison of the altimetry measurement with the tide gauge reading during the satellite passage, before computation and comparison of daily and monthly means, that are obtained at the tide gauges from the times of the satellite passes (Fenoglio-Marc et al., 2012). In this case, a more careful consideration of all corrections applied to altimetry data (including sea state bias, tidal loading and pole tide) is necessary.

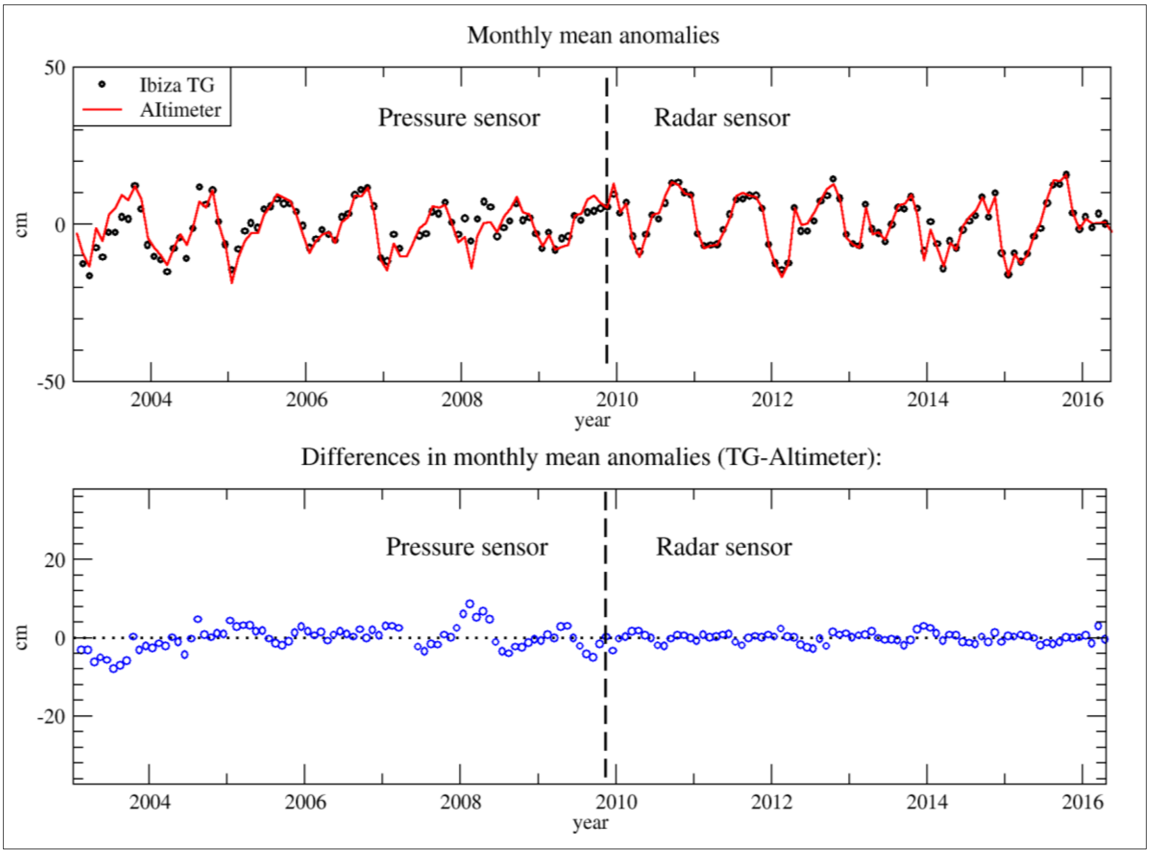


Figure 10. Comparison of monthly means (long-term average removed) between tide gauge and nearby altimetry data at Ibiza tide gauge (Balearic Islands, Spain). Top: monthly means from the tide gauge (black) and from the nearby altimetry data (red). Bottom: differences between the two time series. Vertical black line shows the moment when the radar sensor was installed.

(iii) Datum control

Long-term measurements from a tide gauge provide a record of the height of the sea with respect to the land: a relative sea level height. In order to preserve the vertical reference frame of a sea level series, all heights should be linked to a tide gauge benchmark (TGBM), a physical mark on a stable surface (UNESCO/IOC, 2006). Levelling should be carried out regularly (ideally annually) to a network of nearby benchmarks to ensure that the TGBM is at least locally stable, and to ensure the zero of the tide gauge remains a fixed distance below the height of the TGBM. Wherever possible, a GNSS receiver should be installed at or near the TGBM to provide estimates of vertical land motion, and to allow links to a reference ellipsoid (Wöppelmann et al., 2007; UNESCO/IOC 2016; Woodworth et al., 2017a).

From the PSMSL's perspective, the general principle of datum control for tide gauges is that all heights are referred to the TGBM. Therefore, the series will contain not just information about changes in sea level height, but also changes in the height of the land itself, including long-term changes such as Glacial Isostatic Adjustment, sudden changes from earthquakes, or anthropogenic changes, such as subsidence due to groundwater extraction. The series will not be adjusted as a result of such effects, unless they cause the height difference between the TGBM and the tide gauge itself to change.

(iv) Comparison with other available information depending on the characteristics of the station

As stated above, problems in tide gauge records depend strongly on the characteristics of the station, from the location to the type of instrument, and on the conditions of the surrounding environment. Therefore, an ultimate detailed quality control of a long record will be required to have all potential contributions in mind, and compilation of all the available external information and metadata that can be gathered at each individual site. For example, to understand the influence of the technology or type of instrument, as already stated in all GLOSS manuals (UNESCO/IOC, 1985, 1994, 2002, 2006, 2016), the following should be taken into account:

- Pressure sensors: the influence of errors due to an incorrect assumed density (a clear example has been shown in Figure 8). Water temperature and salinity data could help to correct these effects, when available. On the other hand, when using B-gauges (additional pressure sensors located at a fixed point around mean sea level, for datum check), the analyst should use software to ensure consistency with the B-datum.

- Radar gauges: exact position of the Point of Zero Range (PZR) and Sensor Offset (SO) should be well known and documented, as well as the regular calibration information ideally performed periodically in the laboratory (UNESCO/IOC, 2016). Also, an object or small boat located below the radar antenna can provide invalid measurements for a certain period of time.

- Float gauges can be affected when the stilling well is clogged up.

Most tide gauge records are not corrected for land movement, so unusual changes observed on the reference can be verified with data from a nearby GNSS station that will contain the information on land movement. Use of air pressure, wind and even wave local data can also be of help to understand specific signals during extreme events, or longer term effects in monthly mean sea levels. Many other sources of information can be used: sudden jumps can be compared to times of earthquakes; interannual variability can be compared with known global circulation indices (ENSO, NAO, etc.); short-term spikes can be compared with reports of flooding. Any such events should be included in station documentation to aid interpretation of the data by future users.

2.6 EXISTING SOFTWARE PACKAGES FOR DELAYED MODE QUALITY CONTROL AND PROCESSING

The software packages for delayed mode quality control, available to share with national agencies or scientists that do not have their own code are described in UNESCO/IOC (2016), Section 8.2.3. These include software for tidal analysis and prediction and computation of daily and monthly means. An updated list is provided below:

- The Joint Archive for Sea Level (JASL)/UHSLC package:
  https://www.ncddc.noaa.gov/activities/climate/jasl/index.html

- The NOC (UK) NOCTide packages:
  Included in code at https://psmsl.org/cme/autoqc.php

- The IOS Tidal Package, created by Mike Foreman:
  https://www.dfo-mpo.gc.ca/science/data-donnees/tidal-marees/index-eng.html

- Rich Pawlowicz's t_tide Matlab package:
  https://www.eoas.ubc.ca/~rich/#T_Tide

- Daniel Codiga's UTide Matlab package: https://www.mathworks.com/matlabcentral/fileexchange/46523-utide-unified-tidal-analysis-and-prediction-functions
- A Python implementation of UTide: https://github.com/wesleybowman/UTide

2.7 QUALITY CONTROL FLAGS

Tide gauge datasets should be composed, whenever possible, of the processed data provided by the data collecting authority (i.e. wrong values already eliminated and/or interpolated), with the exception of real time data provided as part of tsunami monitoring systems. The quality controlled data are used for various applications, including providing the data to national, regional and global data centres. Thus, after real time and delayed mode quality control procedures, extensive use of flags to indicate the data quality is vital since the end user will select data based on quality flags amongst other criteria. These flags always need to be included with any data transfer that takes place to maintain standards and to ensure data consistency and reliability.

Table 2 shows the recommended single character qualifying flags which may be associated with one or more individual parameters within a sea level time series. It is important to note that from this scheme, the codes 1 (correct value), 4 (bad data) and 9 (missing value) are mandatory for the near-real time quality control procedures. If no quality control has been carried out, data values should be labelled with code 0 (no quality control). This list of quality control flags has been derived from internationally agreed quality flag schemes (as used by global projects, e.g. Argo, Global Temperature and Salinity Profile Project (GTSPP), etc.), GESLA and several European projects: CMEMS, SeadataNet and European Sea Level Service Research Infrastructure (ESEAS-RI). The quality flags in Table 2 are consistent across all of these schemes. However, there are also additional quality flags included which have different meanings in the projects and programmes above; these are not included in the table. For example, quality flag 2 is defined as "interpolated value" for GESLA and ESEAS-RI, but as "probably good" in the other schemes, some of which use quality flag 8 for "interpolated value". Because of this difference, quality flags 2 and 8 are not included in Table 2. All except GESLA use 9 for "missing value" and this is used in Table 2; GESLA uses quality flag 5 for this. So to improve interoperability and the "FAIR-ness" of the data, the quality flags in Table 2 are recommended. However, if other additional quality flags are required these should be clearly defined.

The user should take into account that data with a quality control flag = 0 should not be used unless they perform detailed and careful quality control from scratch, and that only measurements with flag = 1 can be used safely without further analyses. Measurements with quality control flag = 4 should be rejected. Finally, quality control flag = 3 indicates that the data are not usable but the data centre has some hope to be able to correct them in delayed mode.

Flag	Meaning	Definition
0	No quality control	No quality control procedures have been applied to the data value.
1	Good	Good quality data value that is not part of any identified malfunction and has been verified as consistent with real phenomena during the quality control process.
3	Probably bad (previously 'doubtful')	Data value recognised as unusual during quality control that forms part of a feature that is probably inconsistent with real phenomena.
4	Bad (previously isolate spike or wrong value)	An obviously erroneous data value.
9	Missing	The data value is missing.

Table 2. Recommended numerical quality flag values for GLOSS.

3. METADATA AND DATA FORMAT

Alongside the tide gauge data themselves, as for any other data, additional information (metadata) is needed not only for quality control and archiving, but also for exchanging data or integrating them into regional or global datasets. In addition, accurate metadata plays an important role in ensuring that there is greater alignment with the 'FAIR' Guiding Principles. Thus good quality metadata contributes to making tide gauge data Findable, Accessible, Interoperable, and Reusable. Agreement on standard data formats that allow easier exchange of data and metadata is an inherent component of this objective.

3.1 TYPE OF METADATA AND APPLICATIONS

The International Organization for Standardization (ISO) has published its standard for discovery metadata (ISO19115 for geo-spatially referenced data) and tide gauge datasets should be described in compliance with this standard. For example, discovery metadata descriptions for tide gauge datasets can be provided to the GEOSS Portal (http://www.geoportal.org), Global Change Master Directory (GCMD, http://gcmd.nasa.gov), or for European sea level datasets: European Directory of Marine Environmental Data Sets (EDMED, http://edmed.seadatanet.org/), or Copernicus Marine Environment Monitoring Service (CMEMS, https://marine.copernicus.eu/)

However, this level of metadata is primarily intended for the discovery of datasets and more detailed metadata are required at every stage of the process from initial data collection, real time or near-real time data transmission, automatic and scientific quality control, to long-term stewardship of the data. Table 3 shows a subset of the metadata required to be stored alongside the tide gauge data (e.g. included in the data file or in a linked database), as agreed by the GLOSS data centres in 2010, with input from national authorities. Although not complete (for example, details on the nearby GNSS receiver, when available, and some other aspects related to operational or real time applications are not included here), this table is a first step for a comprehensive list of agreed metadata that should accompany a tide gauge time series. Not all the required metadata listed here will be attached to a time series, but links to where these data can be downloaded from, reports on data processing, etc., is still possible.

Agency Affiliations		
Field	**Description**	**Example**
Originator	Agency responsible for data collection (e.g. EDMO codes: https://edmo.seadatanet.org/search)	Port Authority of Auckland
Contributor	Agency that provided data to international data centre(s)	Land Information New Zealand
Other	If any other agency supported the creation of the final data (repeatable)	University of Waikato

Maintenance History		
Field	**Description**	**Example**
Date-Time	Date-time of update (ISO 8601 format)	2007-07-07T10:00
Person	Who made update	E. Bradshaw
Agency	Of which agency	BODC
Action	What was done	Update coordinates

Site Location Information		
Field	**Description**	**Example**
Country	Country where station is located (ideally use ISO3166 country code)	USA
Originators Station Name	Name used by agency responsible for the tide station	Galveston, Pier 21
Location Description	Explanation of location (can include maps, diagrams and photographs)	Located on Pier 21 on the Gulf side, mid-way along Galveston Beach
Latitude	Decimal latitude (N+, S-)	29.31666
Longitude	Decimal longitude (E+, W-)	-94.80000
Position Source	How coordinates were determined	GNSS, eye-balled off map, unconfirmed
Precision of Position	Position precision in meters	5 m
Horizontal Datum	Geodetic reference in horizontal	WGS84, NAD, etc.
Time Zone and UTC Offset	Local time zone of station and offset from GMT (UTC)	090W, GMT – 6 hr

Benchmark Description		
Field	**Description**	**Example**
Tide Gauge Benchmark (TGBM)	Description of marker and location (include maps and photographs)	USACE disk stamped “8-201” set in sidewalk of NE corner of Harbor House
Bench Mark 1	Description of marker and location (include maps and photographs)	
Bench Mark N	Description of nth marker and location (include maps and photographs)	

Benchmark Specifics		
Field	**Description**	**Example**
Agency Responsible	Who maintains geodetic surveys	NOAA/NOS
Originator ID	ID used by originator	TGBM
Latitude	Decimal latitude (N+, S-)	29.31666
Longitude	Decimal longitude (E+, W-)	-94.80000
Position Source	How coordinates were determined	GNSS, eye-balled off map, unconfirmed
Precision of Position	Position precision in meters	5 m
Field levelling	How often are they surveyed?	Semi-annually

Station Datum (SD) Definition		
Field	**Description**	**Example**
Definition of SD	Specific definition relative to a calculated datum	1.2 m above mean sea level based on epoch 1983-2001
Originator datum name	Definition defined by originator	Same
TGBM to SD	Relative height (m) TGBM above SD	2.643
Epoch of determination	Define epoch	Based on 1983-2001
Accuracy criteria	Accuracy (m)	0.001

Relationships Between Datum/Benchmarks		
Field	**Description**	**Example**
Relationship 1	Describe relationship in vertical between two reference points	Zero of SL data is 1.282 m below TGBM
Relationship N	Repeat as necessary	TGBM is 0.782 m below aux. BM 1

Sea Level Instrumentation Type		
Field	**Description**	**Example**
Sensor 1 Type	General type (float/well, radar) including modifications and possible effect on the data	Radar
	Instrument characteristics	
	Operational history	
Sensor N Type	Repeat for each	Float/well

Sea Level Instrumentation Type Specifics		
Field	**Description**	**Example**
Make	Make of sensor	Sensor Company XXXX
Model	Model of sensor	611-A
Date Installed	When installed (ISO 8601 format)	2001-03-23T00:00
Date Terminated	When removed (blank if not)	

Ancillary Instrumentation Type		
Field	**Description**	**Example**
Sensor 1 Type	General type	barometer
Sensor N Type	Repeat for each	Thermometer for air

Ancillary Instrumentation Type Specifics		
Field	**Description**	**Example**
Parameter	Field measured	Barometric air pressure
Make	Make of sensor	Sensor Company YYY
Model	Model of sensor	QC111
Date Installed	When installed (ISO 8601 format)	2001-03-23T00:00
Date Terminated	When removed (blank if not)	
Originator	Authority in charge of sensor	NOAA/NWS
URL to real time data	Link to data if available	https://tidesandcurrents.noaa.gov/waterlevels.html?id=8724580
URL to historic data	If historic data are available	http://ncdc.noaa.gov

Delayed Mode Processing		
Field	**Description**	**Example**
Calibration Methods	Describe how data calibrated	Tide staff readings and switch data
Data reduction methods	Describe techniques to reduce the sample interval	Hourly data acquired using a 3-pt Hanning
Gap-filling methods	If auxiliary sensor data used to fill gaps, or if data interpolated, describe	Predicted tides method
Quality control methods	Describe editing associated with quality control, or refer to QC document	Timing offsets of exact increments of 1 hour, spikes and short glitches < 24 hours interpolated
Fastest interval	State shortest sample interval of research quality data	Hourly

Delayed Mode Data File Attributes		
Field	**Description**	**Example**
Data units	Scientific units in data files	Millimeters (mm)
Data Time Zone	Offset (hours) relative to UTC	0
Data missing flag	Numeric flag for missing data	99999
Quality/processing flags	Are other flags included with data?	Y
Interpolation	Have short gaps < day been interpolated?	N
Start of Verified Data Date Span	Beginning of time series (ISO 8601 format)	1982-01-01T00:00
End of Verified Data Date Span	End of time series (ISO 8601 format)	2018-12-31T23:00

Table 3. Type of metadata and applications.

Basic relevant metadata include details of the responsible agency, instrumentation and sensors, data sampling, calibration, maintenance operations and data processing, report on quality and a QC history. Other metadata may be required by near-real time operations, for example: latency, communication/transmission channel, date of last received data, status of the station (operational or not), etc. (see the IOC Sea Level Station Monitoring Facility, for further information). Similarly, for long-term studies of mean sea level, detailed metadata from nearby or co-located continuous GNSS measurements is essential to determine absolute mean sea level variations without contamination of vertical land movement and absolute sea levels referred to the ellipsoid. It is strongly recommended to include the ARP (antenna reference point) of a nearby GNSS to the network of benchmarks used to monitor the stability of the tide gauge surrounding. The geodetic tie to the GNSS station should provide: distance between tide gauge and GNSS, ellipsoidal height of the tide gauge benchmark, ellipsoidal height of the tide gauge datum and the trend of vertical land movement at the GNSS station. In GLOSS this metadata information is stored and managed by SONEL (Système d'Observation du Niveau des Eaux Littorales, http://www.sonel.org/), the dedicated GLOSS data centre for GNSS at tide gauges.

Therefore, metadata are also important and required to allow assessment of the usefulness of a time series for a particular application and to ensure that the data are fully documented. It is thus necessary to gather as much relevant information about the tide gauge, data and benchmarks, which could include maps, diagrams and photographs and store them alongside the data. Benchmark history should also be documented.

Moreover, in order to carry out some of the quality control procedures described above, some basic additional information (part of the metadata) must be included for each particular tide gauge station, as input for the quality control procedures. For example, harmonic constants of one year of data (at least 68 constituents), maximum – minimum expected water levels (for out of range detection), or maximum – minimum expected surge values.

Basic quality control of these metadata includes checking that some of them have reasonable values, as is the case for latitude and longitude, the start and end dates of a record, etc. The units employed for each parameter must belong to the Système International (SI). Where available, internationally agreed codes or controlled vocabularies should be used (for example the European Directory of Marine Organisations (EDMO) code for the data provider: https://edmo.seadatanet.org/search is recommended within CMEMS data portals), or other examples as described in https://vocab.nerc.ac.uk.

3.2 RECOMMENDATIONS FOR FUTURE UPGRADE OF METADATA INFORMATION

If metadata are critical, nearly as important is the way the metadata are provided and made available to scientists and users, or for data exchange between different data portals or aggregators. Most of the required tide gauge metadata mentioned above and described in detail in Table 3 have been defined and recommended by GLOSS a decade ago. As already mentioned, more recently, the use of tide gauge data for tsunami warning or near-real time applications has extended the metadata required. Today, operational use of data in near-real time cannot rely just on sporadic manual access to this information on websites, and there has been an increase in machine to machine aggregation of data; therefore “key metadata on provenance and quality can too easily be decoupled from raw data sets, to the detriment of all” (Buck et al., 2019). Part of the metadata can and should be attached to the time series in adequate and agreed standard formats, according to a predefined common definition of “attributes”. This is easily done with formats like NetCDF, which are used by some international data centres within and outside the GLOSS community.

In Europe, the EuroGOOS Tide Gauge Task Team, in collaboration with the GLOSS data centres, has been focusing on this challenge in recent years. Aligned with this objective is the document with recommendations for basic NetCDF attributes of tide gauge data, prepared for CMEMS in May 2017: Recommendations for CMEMS on standard NetCdf format for tide gauge data. Another example is the effort of compiling available information and basic metadata (from SONEL or national institutions) of the European tide gauges co-located and tied to a GNSS station (with the goal of this information being easily included in CMEMS or GLOSS NetCDF data files, public and easy to find by any simple downloading script): http://eurogoos.eu/download/TG_GNSS_2018.pdf. Finally, based on previous initiatives (e.g. IOC Sea Level Station Monitoring Facility), a new metadata portal for European tide gauges is being developed in the region by the Marine Institute of Ireland, in order to integrate all the mentioned relevant information from a particular tide gauge at a single site, to facilitate quick updates from the data providers with visualization tools (maps) of available/missing information, and to automatically generate metadata in the adequate and agreed formats. This activity takes place in collaboration with the GLOSS Data Centres, to guarantee metadata standards and quality and to ensure any improvements to metadata access in Europe are compatible with any subsequent GLOSS implementation in the region.

The OceanObs19 conference (16–20 September, Hawaii, USA) produced several white papers with specific recommendations that will apply to metadata in the forthcoming decade. Two that could apply to sea level data are:

- "Data and metadata are available via standards-based secured API's, using FAIR principles to define data services, to enable new and existing communities to develop their own bespoke web portals, applications, and value-add systems, based on a single digitally-signed quality-controlled data source, to deliver greater uptake, use and value from the collected data." (Buck et al., 2019).

- "Data sets, models and data products are uniquely identified using Digital Object Identifiers (DOI's), digitally signed using certificates to identify source and provenance (including identifying the definitive version of a data set), quality controlled using documented best practice systems (including Quality Control as a Service—QCaaS) with the QC data traveling with or linked to the source data, full machine readable metadata available that includes appropriate use and attribution, as source components of new work-flows." (Buck et al., 2019).

3.3 RECOMMENDATIONS FOR TIDE GAUGE DATA FORMATS

GLOSS has not defined a specific format for storage or exchange of tide gauge data. However, GLOSS recommends that good practice is followed. In this context, tide gauge data should be supplied in a fully documented ASCII format or in Unidata's NetCDF (https://www.unidata.ucar.edu/software/netcdf/). Individual fields, units, etc. should be clearly defined and time zone stated. Time reported in UTC is strongly recommended.

As far as possible, formats should be self-describing and interoperable (i.e. using controlled vocabularies). For this reason, for NetCDF files, it is strongly recommended to use the CF (Climate and Forecast) standard (http://cfconventions.org/), and the Attribute Convention for Data Discovery (ACDD, http://wiki.esipfed.org/index.php/Attribute_Convention_for_Data_Discovery). The CF standard implements a controlled vocabulary through the standard name attribute of variables: these should come from the list at http://cfconventions.org/standard-names.html.

Examples of tide gauge data formats are:

1. **ASCII format** – an example is the columnar format like that used by GESLA (see: https://www.gesla.org/sea_level_format_description_v4.0.pdf). In addition, UHSLC have a long established ASCII format for their hourly and daily products (see: ftp://ftp.soest.hawaii.edu/uhslc/rqds, filename: hourly_format.txt).

2. **CF compliant NetCDF** – NOAA have developed a set of templates (see: https://www.nodc.noaa.gov/data/formats/netcdf/v2.0/). The UHSLC have adopted one of these to distribute their datasets. In Europe CMEMS distributes in CF compliant NetCDF. This is described in their CMEMS NetCDF manual: https://archimer.ifremer.fr/doc/00488/59938/63148.pdf.

Some example file extracts for the formats described above (CDLs for NetCDF) are provided in Annex II.

In addition, as noted in the previous section, the EuroGOOS Tide Gauge Task Team have provided some recommendations for improving the CMEMS NetCDF for tide gauge data, including variable names, mandatory and optional attributes to be included in the file: http://eurogoos.eu/download/NetCdf_Recommendations_forCMEMS_EuroGOOSTGTT_October_2017.pdf.

4. SOFTWARE PACKAGES FOR AUTOMATIC SEA LEVEL DATA QUALITY CONTROL AND PROCESSING

Many institutions have followed most of the recommended steps of quality control and data processing described in Section 2. However, not all of them have an automatic code that is applied in near-real time (Section 2.3.1) or that, applied in delayed mode, can reduce the degree of human intervention to a minimum. In some cases, the automatic processes exist but are based on old code or are still under development, so they are difficult to share with the rest of the sea level community. To our knowledge, and at the time of writing, the following institutions already have some automatic quality control of tide gauge data in place:

- Australia: The Australian Bureau of Meteorology (BoM) performs automatic real time quality control by comparing predictions with observations and some elementary range checking. This method is applied to 1-min data. Contact: Bill Mitchell (bill.mitchell@bom.gov.au)

- Canada: The Canadian Hydrographic Service (DFO-CHS) is developing national standardized primary real time automated QC followed by secondary manual QC to apply corrections as needed to the data for final archival. Primarily developed to service emerging Dynamic E-navigational requirements. The observational data will be combined with ocean model forecasts to provide S104 (water level) and IHO S111 (surface current) products. Their primary automated QC processes are based on inter-comparisons between 3 sensors at each gauge location coupled with comparison with forecast data. They are now working on enhancements, including additions for additional QC capacities. The project and capacities are still under development. Contact: Philip MacAulay (Phillip.MacAulay@dfo-mpo.gc.ca)

- Italy: The Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA) carries out real time quality control based on in-house developed software. In addition, delayed mode semi-automatic quality control is carried out on monthly and yearly time series. Data are validated according to residual analysis, comparison between contiguous instruments and expert judgement. Contact: Marco Picone (rete.mareografica@isprambiente.it)

- Mexico: The Servicio Mareográfico Nacional of the Instituto de Geofísica, Universidad Nacional Autónoma de México (UNAM), is currently developing an automated system to carry out L1 QC for 1-min data that is published in near real time on the web portal and distributed among their partners, as well as automated tools to assist technicians who carry out L2 QC. The system is being developed considering, among others, the recommendations for GLOSS stations. Once the system is completed, it will seek certification under ISO 9001: 2015, and the Python source code will be publicly available. The development of the project continues and is expected to be completed by the end of 2020. Contact: Octavio Gómez (octavio@geofisica.unam.mx)

- New Zealand: The National Institute of Water and Atmospheric Research (NIWA) has developed semi-automatic quality control software based on a suite of code that runs in MATLAB and hinges on the non-tidal residual (NTR). The key steps the code targets are: de-spiking relative to the standard deviation or an absolute value of NTR (excluding known tsunami events), check timing (including daylight-saving time), treatment of gaps and glitches, with optional interpolation (only up to 24 hours) and checks for datum shifts. Contact: Scott Stephens (scott.stephens@niwa.co.nz)

- Norway: Kartverket uses an automatic quality control developed in-house for the Norwegian tide gauge network. Contact: Oda Roaldsdotter Ravndal (oda.ravndal@kartverket.no)

- Republic of Korea: Korea Hydrographic and Oceanographic Agency (KHOA) performs quality control in two steps. In the first, an automatic process flags data based on the variance evolution of the time series. In the second, manual quality control is performed by data managers who consider other ocean environmental factors, comparison with nearby stations, historical data and predictions. Contact: Aram Kim (1124kar@korea.kr)

- South Africa: South African Navy Hydrographic Office (SANHO) performs a bi-weekly manual, visual check by graphically comparing the predictions with the observations (either 1 or 3 minute observations). The annual raw data set is converted to hourly data that is processed on a yearly basis through the SANHO's in-house software. Using a Lagrange polynomial fit, the software identifies and "flags" any anomalous data or deviations for further investigation against the predictions and bi-weekly graphics. In 2010, an additional automated algorithm was added that fits a 3rd order polynomial curve through the "flagged" data, this approach provides an extra quality assurance step. The "flagged" data can be replaced by the calculated value or removed all together. Contact: Ruth Farre (hydrosan@iafrica.com)

- Spain: Puertos del Estado has performed near-real time automatic quality control of their tide gauge network REDMAR since 1998. The method is the one described in this document and later on adopted by ESEAS-RI project, EuroGOOS and CMEMS. Since 2018 a new Python version (SELENE) has been prepared for sharing with other institutions and experts, via a Github project: https://puertos-del-estado-medio-fisico.github.io/SELENE/. Documentation: https://github.com/puertos-del-estado-medio-fisico/SELENE/blob/master/documentation/SELENE_DesignAndUsersGuide.pdf Contact: Begoña Pérez Gómez (bego@puertos.es)

- Sweden: The Swedish Maritime Administration (SMA) has also implemented the near-real time QC procedures described here and recommended by CMEMS, EuroGOOS, IODE and QARTOD, to the Swedish Sea Level network (60 stations). Contact: Thomas Hammarklint (thomas.hammarklint@sjofartsverket.se)

- The Netherlands: At Rijkswaterstaat automated quality control is conducted on sensor signals, samples and calculated values: checks on sensor operation, checks on the used samples (like number of samples, delta check, min/max check etc.), value checks (like reference checks against values of nearby locations or back-up sensors). Also, semi-automated quality control is performed on calculated values by flagging deviations (M2 tide control) via visual inspection of data. All controls are location specific and based on historic behaviour. (Email: helpdeskwater@rws.nl)

- United Kingdom: The National Oceanography Centre (NOC) has developed automatic quality control software as part of a UK Government effort to develop regional capacity building in small island developing states. Available from the PSMSL website at https://psmsl.org/cme. The code runs in MATLAB, and includes a MATLAB implementation of the in-house tidal analysis software developed over many years by NOC and its predecessor organizations. A more detailed description is in Williams et al., 2019. Contact: Andrew Matthews (antt@noc.ac.uk).

- United States: NOAA, as part of the QARTOD (Quality Assurance of Real Time Oceanographic Data) group, already mentioned, has been performing automatic real time quality control to their tide gauge data for several years (https://cdn.ioos.noaa.gov/media/2017/12/qartod_water_level_manual.pdf). Contact: Peter Stone (peter.stone@noaa.gov)

4.1 TESTING AUTOMATIC QUALITY CONTROL AND PROCESSING SOFTWARE: METRICS AND RECOMMENDATIONS

Any automation process must be verified before being extensively used on a routine basis, and should approach as far as possible what a qualified expert on sea level data quality control would do. If the automated process is to be implemented in real (or near-real) time the limitations will be larger than when used in delayed mode in combination with more metadata and ancillary information. In any case, these limitations should be well known and the skills of the software well defined. In order to test the skills of automatic quality control software packages, being developed in-house or not, several metrics should therefore be defined, to evaluate if the performance is reasonable for operational applications, which will be the main objective of automation, or to reduce human intervention in delayed mode quality control of long time series.

Focusing here on the product quality of the software output, the most immediate approach is to define which specific errors we expect to find in the data and to be solved by an automatic process. On the other hand, it must be established which real events or situations should never be flagged or considered as wrong values by the software. The latter is particularly important and challenging when dealing with extreme sea level events. One should have in mind the following main limitations for the adequate implementation of this type of software:

- The availability of enough metadata information (see Section 3) that the software can automatically use for the quality control assessment. Today relevant metadata are far from being automatically ready and included in the time series.
- The availability of a secondary sensor or sea level channel that will allow the use of, for example, the "buddy" checking technique.
- Situations in the data, which would also cause a human to be in doubt.
- The sampling of the original time series, e.g. higher-frequency data with 1 min or less sampling are more challenging than monthly means due to the larger variability.
- The precision of the original data: low-precision data (e.g. cm instead of mm) can be more challenging, particularly in combination with high-frequency sampling.
- The length of the time series: if sufficiently long, tide and non-tidal residuals can be computed and included in the process, and datum changes identified in monthly mean time series.
- The implementation in real time, near-real time or delayed mode: as described in Section 2 the possibilities for real time quality control are very limited, if any. Assessing the quality of each individual data point will require, for most algorithms, a certain temporal window of data centered on that particular measurement (e.g. the spline algorithm described in Section 2).

According to this, one may first establish which errors are easier to detect automatically, based on the experience gained by the institutions already using this kind of approach. This will be the minimum required for the software package to be useful in near-real time operations, and could be of help as well for accomplishing delayed mode quality control. It should be assumed, however, that these useful tools, due to the reasons above, are not perfect and sufficient for all possible situations and types of problems. Existing code will certainly evolve in the future and new possibilities will arise thanks to the use of new techniques such as machine learning and artificial intelligence, which may expand the number of problems automatically fixed in the data in the coming years, and as more standardized and complete metadata are included in the process. The methodology for testing the software suggested here is therefore a first step that hopefully will lead to new projects and exploration of new techniques in the future.

As described in Section 2, these are the different types of well-known errors that can be found in tide gauge data (most of them common to any oceanographic/meteorological data):

- constant values,
- near-constant values
- out of range values (this is dependent on the limits defined for the station),
- isolated spikes,
- clustered spikes,
- datum changes,
- drifts,
- time errors.

At the time of writing, most of the existing automatic software packages do not deal or correct the last three (datum changes, drifts and time errors), which is understandable because they require more involvement of sea level expertise and ancillary information.

These are, on the other hand, the most challenging real situations when automatic software packages should not flag real data as an error:

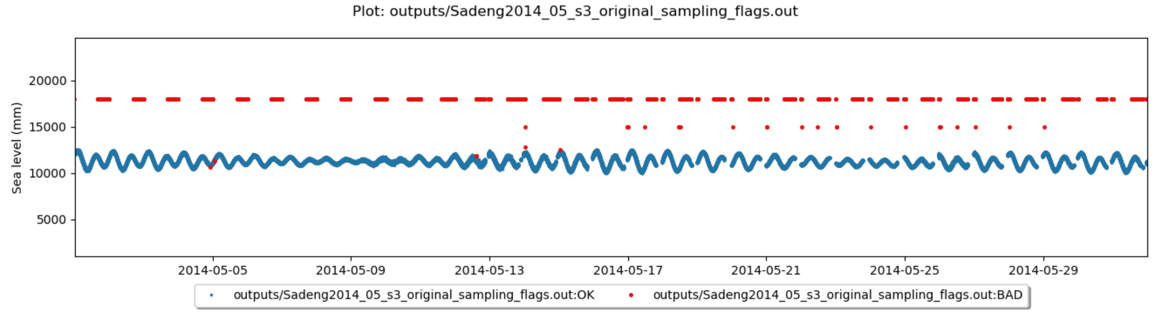
- tsunami/meteotsunami events,
- extreme sea levels during a storm surge.

Therefore, our recommendation is to select for testing examples of two types of sea level time series: those that present very clear wrong data, expected to be flagged by the software, according to its characteristics and capabilities, and those that may be challenging because they contain real events that the software should be able to identify as real.

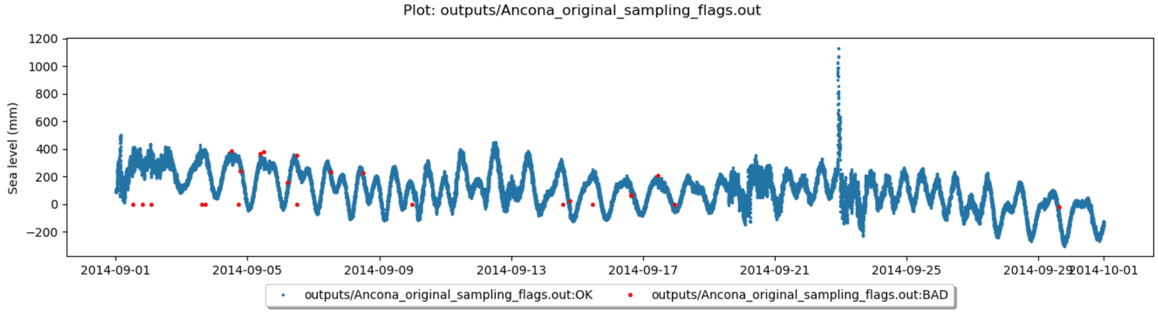
Figure 11 shows some examples of both types of time series, as obtained, for example, from applying SELENE quality control software. Wrong data in this and the following figures are displayed in red, and correct data in blue. It can be seen that wrong data are well identified and flagged for Sadeng tide gauge ([Figure 11a], original time sampling 20s), an example with clear and easy to detect type of errors: constant values or clear out of range values. Figure 11b shows an extreme sea level event recorded at Ancona tide gauge (Italy) on 22 September 2014 (original time sampling 1 min). In this case, the real event is correctly unflagged (although other wrong data are found and flagged adequately). The magnitude of the event was remarkably larger than the local sea level variability (small tides), which makes this event particularly challenging for automatic quality control algorithms. Another challenging situation is shown in Figure 11c, for a time series from Meulaboh tide gauge (original sampling 20s) containing a tsunami on April 11, 2012, also correctly identified as a real event.

The characteristics of the software is a relevant aspect here: as already mentioned, according to the information from existing packages, available to the public or not at this moment, none of them are able to identify and correct all the obvious errors listed above. In addition, there may be small differences between different software packages. For example: SELENE software should successfully flag constant values, out-of-range values, isolated and clustered spikes, but cannot yet automatically detect datum jumps, drifts and time errors. However, it is able to detect and flag temporary datum changes (usually due to a floating object or small boat below the radar antenna), lasting less than approximately 30 min. Longer persisting problems are more likely to be considered correct by the software. The NOC software performs to a similar standard, but can also identify cases where multiple sensors diverge, although cannot pinpoint the exact time divergence in cases where the change is gradual.

a) Time series with clearly wrong data:



b) Time series with a real event (storm surge), challenging for the quality control:



c) Time series with a real event (tsunami), challenging for the quality control:

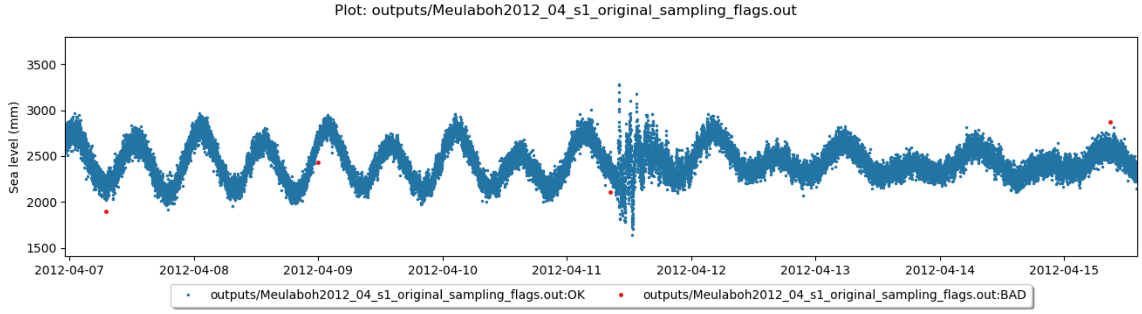


Figure 11. Examples of output of a quality control software for (a) a time series with clear wrong data, including spikes, constant values and out of range data, (b) a time series with an extreme sea level caused by a storm, and (c) a time series with a tsunami event. Flagged (wrong) values in red colour, correct data in blue (details in the text).

In the examples shown above only the first step of quality control as described in Section 2.3.1 was used: the second step based on tide availability and spline-fit applied to non-tidal residuals can improve the performance of the software, as also shown in that section.

4.2 EXAMPLES OF PROBLEMS AND LIMITATIONS OF THE AUTOMATIC QUALITY CONTROL

Understanding the limitations of the software is as important as knowing its capabilities. Some problems will be shown below as an example.

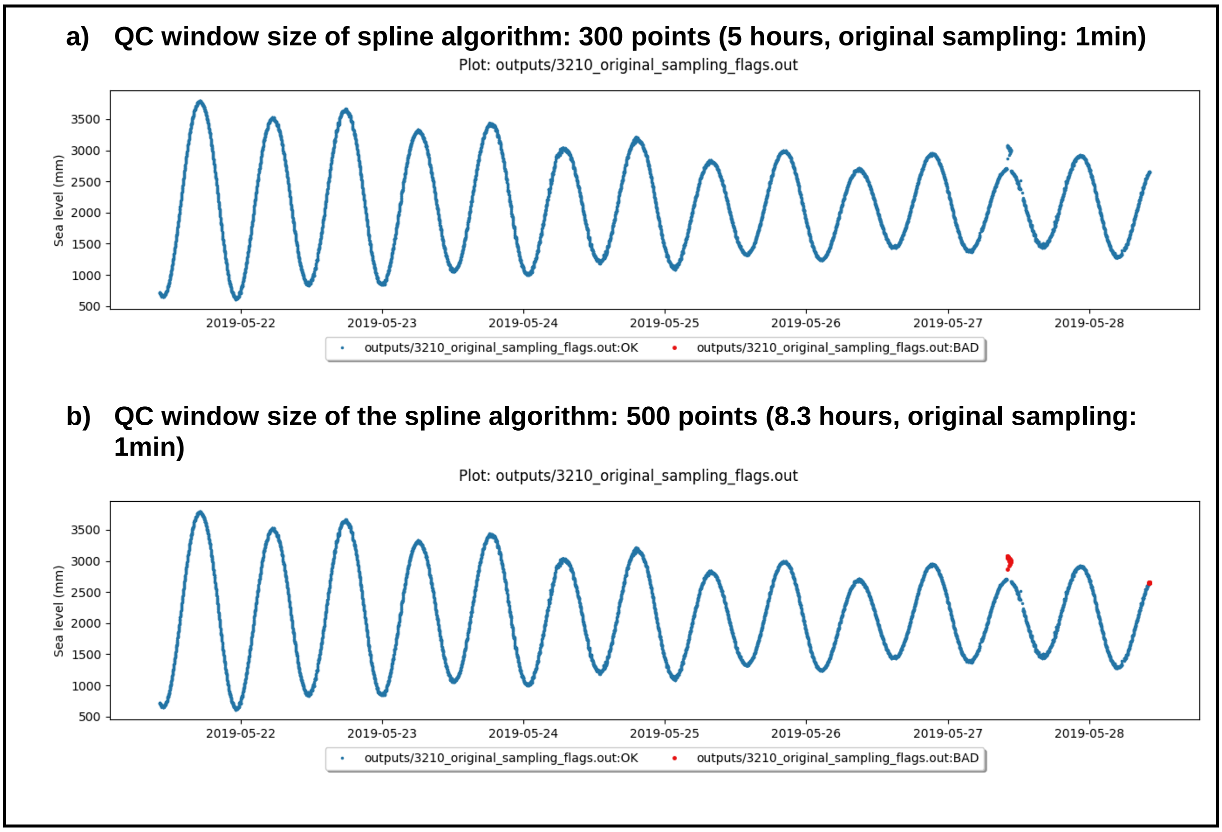


Figure 12. Output of automatic quality control for a radar time series with incorrect data due to an external object below the antenna: (a) the bad data are not flagged if the window size of the qc module (spline algorithm) is 300 points (5 hours, same result for 1 hour); (b) if window size is extended to 500 points (8.3 hours), the data are correctly flagged. Flagged (wrong) values in red colour, correct data in blue (details in the text).

Figure 12 shows the difficulty of flagging not valid data caused by an object located below a radar sensor, and the dependency on the configuration of the spline algorithm. One can consider this as a temporary change of reference or as clouded spikes, always more difficult to process than individual incorrect values. Figure 12a displays the output of the automatic quality control for NOC and SELENE software (similar output), by using a window size of 1 and 5 hours, respectively, for the spline algorithm; in this case the data are not identified as wrong by the algorithm. However, if the window size is increased to 8.3 hours, the event is clearly identified as invalid data (Figure 12b). The longer the object remains below the antenna, the more difficult it will be to get the software working in this situation. Use of the tide (if available) and non-tidal residual in the second step (not shown here) can also help to get these data flagged.

However, the most challenging and problematic aspect is the correct identification of extreme sea levels. This is easy if the event is represented by several points on the time series, as the examples in Figure 11. However, especially if the sampling is not sufficiently small (1 min or less), an extreme sea level event can be represented by a single point, and therefore considered a spike by the existing algorithms described in this manual. An example is presented in Figure 13, for an extreme sea level recorded by NOAA Clearwater Beach tide gauge during Hurricane Michael. Both SELENE and NOC software packages flagged this value as wrong. Only reducing the time window of the algorithm to an hour would prevent the individual data to be flagged. In this case, the data sampling is 6 min.

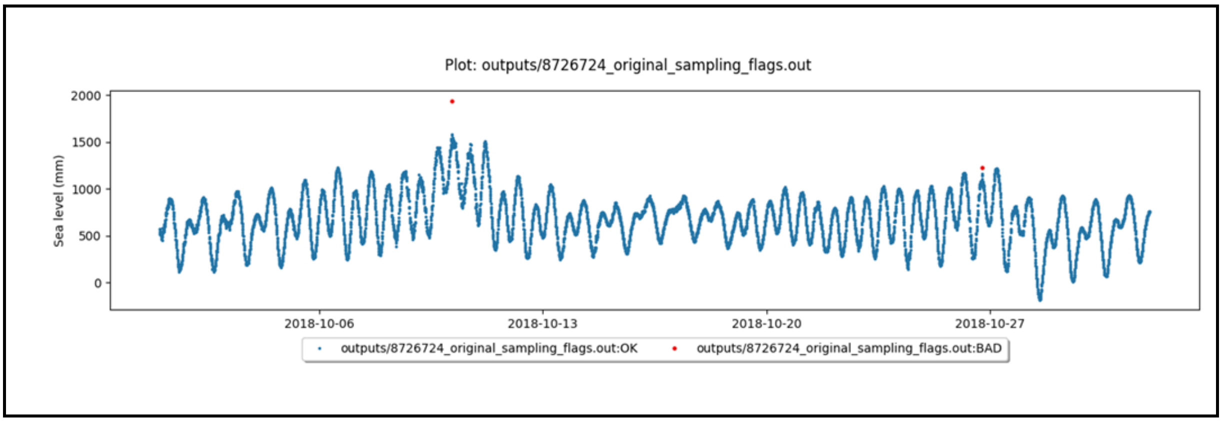


Figure 13. Output of SELENE software for Clearwater Beach: flagged (wrong) values in red colour. The extreme real sea level event is incorrectly flagged as wrong by the software. The 6min sampling of the original time series complicates things because it is too long probably for this kind of event.

4.3 DISCUSSION

The examples presented in this section are representative of the kind of situations faced in the process of quality control, and the difficulties and limitations of existing algorithms. In some cases, manual quality control would be equally challenging and a clear distinction between correct and incorrect data could remain open. This has been traditionally solved with the "doubt" flag, something that none of the software packages shown here include at this moment (one of the possible lines of improvement).

As already mentioned, several institutions are working continuously on the improvement of these automatic procedures and great advantage could be gained in the future from easier access to metadata, ancillary data, nearby sensors and more sophisticated algorithms (machine learning techniques). Some errors are expected to become less common (e.g., as clocks synchronized to GNSS or the Internet may reduce timing errors).

5. CONCLUSIONS

Despite the long history of tide gauge measurements, and the well-known basic steps required for their quality control and processing, this is the first time all relevant aspects of, and recommendations on, this topic have been gathered into one manual. As stated in the introduction, most of the content is not new, but has been compiled from previous international programmes and projects documentation, including other GLOSS manuals. These recommendations have been updated for several reasons:

- New applications of tide gauge data in recent years: the multi-purpose approach has increased significantly the number of data portals and users, posing a challenge for maintaining standards on data quality control and processing.

- New requirements on latency and sampling (< 1min), for tsunami warning: this has introduced differences in the way data are obtained and on the data flow, with a clear impact on the first steps of quality control.

- The increase of operational applications: this is reflected in the convenience of adopting automatic procedures that minimize the impact of invalid data in near-real time applications (e.g. storm surge forecasts, routine model validation and assimilation or altimetry comparison, etc.).

- Technology changes: new types of tide gauges, e.g. radar sensors, are more widespread today, presenting differences on the type of errors most commonly observed.

- The increased volume of high-frequency data available worldwide without quality control.

The need for automatic quality control, in near-real time or in delayed mode, is one of the main challenges. Apart from the reasons detailed above, one of the problems some institutions face is the lack of sufficient human resources for manual and careful inspection of the data. Although this will ultimately be needed, the possibility of reducing the amount of time dedicated to visual inspection will be of great help. As demonstrated several institutions are already applying automatic quality control procedures, or have established the first steps of the quality control, and there will probably be significant efforts in the future to improve the skills of these automatic processes. Hopefully, this may foster collaboration between institutions and experts on, for example, new techniques involving big data and machine learning approaches.

However, whatever is implemented, comprehensive tests should be made. As shown in Section 4, there are situations where even an analyst could have doubts, and this problem affects both extreme sea level events (affected by the data sampling and high sea level variability) and the computation of reliable long-term sea level trends (affected by small reference problems or inhomogeneities that will certainly be difficult to identify).

From the detailed review provided in this manual, several final thoughts and recommendations can be provided. For example, more complete metadata information should be available and easily linked to a tide gauge time series; this will be crucial for developing new tools that could easily exploit these metadata and improve the skills of the automatic software packages (e.g. tidal constituents, dates of sensor replacement or maintenance, land movement from a GNSS, etc.). For this, a great effort of standardization is needed and, compiling (and updating) all relevant information may be demanding and require significant resources. Development of tools that facilitate this work to tide gauge operators and GLOSS Data Centres would be highly desirable. Installation of secondary sensors (as already recommended by GLOSS in previous manuals) will also increase the possibilities of automatic algorithms and the use of the “buddy checking” approach that can be useful for automatic detection and flagging of reference problems. This is not possible today for all tide gauge stations, in spite of it being the standard recommendation for the GLOSS core network. Lastly, cross-comparison of tide gauge data with other relevant parameters or ancillary information in an automatic way could also improve the possibilities of automatic quality control: for example, information on weather conditions (wind, waves or atmospheric pressure data) could be a valuable tool to discard invalid flagging during extreme sea level events.

Finally, the amount of quality control that can be done in near-real time applications will always be limited and will not replace the detailed and more complete analysis required for historical time series.

6. REFERENCES

Alexanderson, H. 1986. A homogeneity test applied to precipitation data. *Journal of Climatology*, Vol. 6, Issue 6, pp. 661–675. (https://doi.org/10.1002/joc.3370060607)

Buck, J.J.H. et al. 2019. Ocean data product integration through innovation-The next level of data interoperability. *Frontiers in Marine Science*, Vol. 6, 32 pp. (https://doi.org/10.3389/fmars.2019.00032)

Codiga, D. 2020. UTide Unified Tidal Analysis and Prediction Functions (https://www.mathworks.com/matlabcentral/fileexchange/46523-utide-unified-tidal-analysis-and-prediction-functions), MATLAB Central File Exchange. Retrieved April 6, 2020. http://www.po.gso.uri.edu/~codiga/utide/utide.htm

Fenoglio-Marc, L. et al. 2012. Sea Level Change and Vertical Motion from Satellite Altimetry, Tide Gauges and GPS in the Indonesian Region. *Marine Geodesy*, 35:sup1, 137–150, doi: 10.1080/01490419.2012.718682

Foreman, M.G.G. 1977. Manual for tidal heights analysis and prediction. *Canadian Pacific Marine Science* Report No. 77–10, 10 pp.

García, M.J. et al. 2007. European Sea Level Monitoring: Implementation of ESEAS Quality Control. In: Tregoning P., Rizos C. (eds) Dynamic Planet. *International Association of Geodesy Symposia*, Vol. 130. Springer, Berlin, Heidelberg. (https://doi.org/10.1007/978-3-540-49350-1_11)

Godin, G. 1972. *The Analysis of Tides*. Liverpool University Press, 264 pp

ICES Data and Information Group (DIG). 2006. ICES Guidelines for Water Level Data. (Compiled August 1999, revised August 2001, revised May 2006). Copenhagen, Denmark, International Council for the Exploration of the Sea (ICES), 7 pp. http://ices.dk/sites/pub/Publication%20Reports/Data%20Guidelines/Data_Guidelines_TWL_v7_revised_2006.pdf (Retrieved 9 April 2020)

Marcos, M et al. 2005. Consistency of long sea level time series in the northern coast of Spain. *Journal of Geophysical Research*, 110, C03008 (doi: 10.1029/2004JC002522)

Martín Míguez, B., Testut, L. and Wöppelmann, G. 2008. The Van de Casteele test revisited: an efficient approach to tide gauge error characterization. *Journal of Atmospheric and Oceanic Technology* (doi: 10.1175/2007JTECHO554.1)

Murray, M.T. 1964. A general method for the analysis of hourly heights of the tide. *International Hydrographic Review*, Vol. 41, pp. 91–101

Pawlowicz, R., Beardsley, B. and Lentz, S. 2002. Classical tidal harmonic analysis including error estimates in MATLAB using T_TIDE. *Computers and Geosciences*, Vol. 28, Issue 8, pp. 929–937. (https://doi.org/10.1016/S0098-3004(02)00013-4)

Pérez Gómez, B. et al. 2013. Use of tide gauge data in operational oceanography and sea level hazard warning systems. *Journal of Operational Oceanography*, Vol. 6, issue2, pp. 1–18, (doi: 10.1080/1755876X.2013.11020147)

Pérez Gómez, B. et al. 2014. Overlapping sea level time series measured using different technologies: an example from the REDMAR Spanish network. *Natural Hazards and Earth System Sciences*, Vol. 14, pp. 589–610 (doi: 10.5194/nhess-14-589-2014)

Pouliquen, S. and the EuroGOOS DATA-MEQ working group. 2011. *EuroGOOS: Recommendations for in-situ data Real Time Quality Control*. EuroGOOS Publication Number 27. (http://eurogoos.eu/download/reference_documents_/rtqc.pdf, Retrieved 9 April 2020)

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

Reverdin, G. et al. 2017. *Recommendations for an automatic RTor NRTQC for selected EOVs (T&S, Current, Oxygen, CHla, Nitrate, Carbon, Sea level)*. Deliverable 7.2 from EU funded Horizon2020 AtlantOS project (Grant No. 633211). (https://www.atlantos-h2020.eu/download/7.2-QC-Report.pdf, retrieved 5 April 2020)

Rickards, L.J. and Kilonsky, B.J. 1998. *Developments in Sea Level Data Management and Exchange. Proceedings of the IOC/NOAA/EU-MAST Symposium on Ocean Data for Scientists*. Marine Institute, Dublin, October 1997, 11 pp. plus tables and figures.

Schöne, T., Schön, N. and Thaller, D. 2009. IGS Tide Gauge Benchmark Monitoring Pilot Project (TIGA): Scientific Benefits. *Journal of Geodesy*, Vol. 83, pp. 249–261, (10.1007/s00190-008-0269-y)

SeaDataNet. 2010. *Data Quality Control Procedures*. Version 2.0. May 2010. Deliverable from EU funded FP7 SeaDataNet project (Grant Agreement No. 283607) (https://www.seadatanet.org/content/download/596/file/SeaDataNet_QC_procedures_V2_%28May_2010%29.pdf, Retrieved 9 April 2020).

U.S. Integrated Ocean Observing System (IOOS), QARTOD, 2016. *Manual for Real-Time Quality Control of Water Level Data Version 2.0: A Guide to Quality Control and Quality Assurance of Water Level Observations*. 46 pp

UNESCO/IOC. 1985. *Manual on Sea Level Measurement and Interpretation. Volume I: Basic Procedures.* (eds. Ainscow, B., Blackman, D., Kerridge, J., Pugh, D. and Shaw, S.). Paris, UNESCO, IOC Manuals and Guides No. 14. https://unesdoc.unesco.org/ark:/48223/pf0000065061

UNESCO/IOC. 1993. *Manual of Quality Control Procedures for Validation of Oceanographic Data*. Paris, UNESCO, Manual and Guides no 26. https://unesdoc.unesco.org/ark:/48223/pf0000138825.locale=en

UNESCO/IOC. 1994. *Manual on Sea-Level Measurement and Interpretation. Volume II: Emerging Technologies.* (ed. Blackman, D.). Paris, UNESCO, IOC Manuals and Guides No. 14. https://unesdoc.unesco.org/ark:/48223/pf0000149528

UNESCO/IOC. 1997. *Global Sea Level Observing System (GLOSS) Implementation Plan 1997.* Paris, UNESCO, IOC Technical Series No. 50. https://unesdoc.unesco.org/ark:/48223/pf0000112650

UNESCO/IOC. 2002. *Manual on Sea Level Measurement and Interpretation. Volume III: Reappraisals and recommendations as of the year 2000*. (eds. Woodworth, P. and Aarup, T.). Paris, UNESCO, IOC Manuals and Guides No. 14. https://unesdoc.unesco.org/ark:/48223/pf0000125129.locale=en

UNESCO/IOC. 2006. *Manual on sea-level measurement and interpretation. Volume IV: An Update to 2006*. (eds. Aarup, T., Merrifield, M., Pérez Gómez, B., Vassie, I. and Woodworth, P.). Paris, UNESCO, IOC Manuals and Guides No. 14. https://unesdoc.unesco.org/ark:/48223/pf0000147773

UNESCO/IOC. 2012. *The Global Sea Level Observing System (GLOSS): implementation plan, 2012.* Paris, UNESCO, IOC Technical Series No. 100. https://unesdoc.unesco.org/ark:/48223/pf0000217832

UNESCO/IOC. 2016*. Manual on Sea-level Measurements and Interpretation, Volume V: Radar Gauges.* (eds. Woodworth, P., Aarup, T., André, G., Donato, V., Enet, S., Edwing, R., Heitsenrether, R., Farre, R., Fierro, J., Gaete, J., Foden, P., Pérez Gómez, B., Rickards, L. and Schöne, T.). Paris, UNESCO, IOC Manuals and Guides No. 14. (https://unesdoc.unesco.org/ark:/48223/pf0000246981)

Vinogradov, S.V. and Ponte, R.M. 2011. Low frequency variability in coastal sea level from tide gauges and altimetry. *Journal of Geophysical Research*, Vol. 116, Issue C07 (doi: 10.1029/2011JC007034)

Wilkinson, M.D. et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, Vol. 3, article number 160018 (doi: 10.1038/sdata.2016.18)

Williams, J., Matthews, A. and Jevrejeva, S. 2019. Development of an automatic tide-gauge processing system. Southampton, National Oceanography Centre, 26 pp. (National Oceanography Centre Research and Consultancy Report, 64). (http://nora.nerc.ac.uk/id/eprint/523262/)

Woodworth, P.L. 2017. Differences between mean tide level and mean sea level. *Journal of Geodesy*, Vol. 91, pp. 69–90. (https://doi.org/10.1007/s00190-016-0938-1)

Woodworth, P.L. et al. 2017a. Why we must tie satellite positioning to tide gauge data. *Eos*, Vol. 98, Issue 4. (doi: 10.1029/2017EO064037)

Woodworth, P.L. et al. 2017b Towards a global higher-frequency sea level dataset. *Geoscience Data Journal*, Vol. 3, Issue 2, pp. 50–59 (doi: 10.1002/gdj3.42)

Woodworth, P.L., Pugh, D.T. and Plater, A.J. 2015. Sea level measurements from tide gauges. Chapter 35 (pp.557–574) in *Handbook of Sea-Level Research*, First Edition. Edited by Ian Shennan, Antony J. Long, and Benjamin P. Horton. Published 2015 by John Wiley & Sons, Ltd.

Wöppelmann, G., Aarup, T. and Schöne, T. 2007. *An inventory of collocated and nearly-collocated CGPS stations at tide gauges*. Progress report (as of July 25, 2007) on the survey carried out under the GLOSS program of IOC/UNESCO. The HTML version of this report provides the links to the text file tables of the report.

Wöppelmann, G. and Marcos, M. 2016. Vertical land motion as a key to understanding sea level change and variability. *Reviews of Geophysics*, Vol. 54, Issue 1, pp. 64–92, (doi: 10.1002/2015RG000502)

ANNEX I

IOC OCEANOGRAPHIC DATA EXCHANGE POLICY

Preamble The timely, free and unrestricted international exchange of oceanographic data is essential for the efficient acquisition, integration and use of ocean observations gathered by the countries of the world for a wide variety of purposes including the prediction of weather and climate, the operational forecasting of the marine environment, the preservation of life, the mitigation of human-induced changes in the marine and coastal environment, as well as for the advancement of scientific understanding that makes this possible.

Recognizing the vital importance of these purposes to all humankind and the role of IOC and its programmes in this regard, the Member States of the Intergovernmental Oceanographic Commission agree that the following clauses shall frame the IOC policy for the international exchange of oceanographic data and its associated metadata.

Clause 1: Member States shall provide timely, free and unrestricted access to all data, associated metadata and products generated under the auspices of IOC programmes.

Clause 2: Member States are encouraged to provide timely, free and unrestricted access to relevant data and associated metadata from non-IOC programmes that are essential for application to the preservation of life, beneficial public use and protection of the ocean environment, the forecasting of weather, the operational forecasting of the marine environment, the monitoring and modelling of climate and sustainable development in the marine environment.

Clause 3: Member States are encouraged to provide timely, free and unrestricted access to oceanographic data and associated metadata, as referred to in Clauses 1 and 2 above, for non-commercial use by the research and education communities, provided that any products or results of such use shall be published in the open literature without delay or restriction.

Clause 4: With the objective of encouraging the participation of governmental and non-governmental marine data gathering bodies in international oceanographic data exchange and maximizing the contribution of oceanographic data from all sources, this Policy acknowledges the right of Member States and data originators to determine the terms of such exchange, in a manner consistent with international conventions, where applicable.

Clause 5: Member States shall, to the best practicable degree, use data centres linked to IODE's NODC and WDC network as long-term repositories for oceanographic data and associated metadata. IOC programmes will co-operate with data contributors to ensure that data can be accepted into the appropriate systems and can meet quality requirements.

Clause 6: Member States shall enhance the capacity in developing countries to obtain and manage oceanographic data and information and assist them to benefit fully from the exchange of oceanographic data, associated metadata and products. This shall be achieved through the non-discriminatory transfer of technology and knowledge using appropriate means, including IOC's Training Education and Mutual Assistance (TEMA) programme and through other relevant IOC programmes.

Definitions

‘Free and unrestricted’ means non-discriminatory and without charge. “Without charge”, in the context of this resolution means at no more than the cost of reproduction and delivery, without charge for the data and products themselves.

‘Data’ consists of oceanographic observation data, derived data and gridded fields.

‘Metadata’ is 'data about data' describing the content, quality, condition, and other characteristics of data.

‘Non-commercial’ means not conducted for profit, cost-recovery or re-sale.

‘Timely’ in this context means the distribution of data and/or products, sufficiently rapidly to be of value for a given application

‘Product’ means a value-added enhancement of data applied to a particular application.

ANNEX II

EXAMPLE EXTRACTS FROM SEA LEVEL DATA ASCII AND NETCDF FILES[1]

1. Example of an extract of GESLA data file

```
# FORMAT VERSION 3.0 Web: http://gesla.org Email: gesla.help@gmail.com
# SITE NAME ABASHIRI
# COUNTRY Japan
# CONTRIBUTOR Japan Meteorological Agency
# LATITUDE      44.0167
# LONGITUDE    144.2833
# COORDINATE SYSTEM Unspecified
# START DATE/TIME 1960/12/31 15:00:00
# END DATE/TIME 2013/12/31 14:00:00
# TIME ZONE HOURS 0
# DATUM INFORMATION: Zero of Tide Height
# INSTRUMENT Unspecified
# PRECISION 0.01 (m)
# NULL VALUE -99.9999
#
# CREATION DATE UTC 2016/03/16
#
# COLUMN 1 Date yyyy/mm/dd
# COLUMN 2 Time hh:mm:ss
# COLUMN 3 Observed sea level (m)
# COLUMN 4 Observed sea level QC flag
# COLUMN 5 used-in-extremes-analysis flag (1 = used, 0 = not used)
#
# Quality-control (QC) flags
#
# 0 - no quality control
# 1 - correct value
# 2 - interpolated value
# 3 - doubtful value
# 4 - isolated spike or wrong value
# 5 - missing value
#
1960/12/31 15:00:00           1.1900 1 1
1960/12/31 16:00:00           1.3000 1 1
1960/12/31 17:00:00           1.4000 1 1
1960/12/31 18:00:00           1.4500 1 1
1960/12/31 19:00:00           1.4600 1 1
```

[1] This annex shows extracts of sea level data in two different formats – ASCII as used in the GESLA data set, and NetCDF as used in UHSLC data set. Please note that the format is presented with a “Courrier” font and text that does not fit in 1 line will automatically wrap to next line.

```
1960/12/31 20:00:00           1.4400 1 1
1960/12/31 21:00:00           1.4100 1 1
1960/12/31 22:00:00           1.4200 1 1
1960/12/31 23:00:00           1.4800 1 1
1961/01/01 00:00:00           1.5700 1 1
1961/01/01 01:00:00           1.6600 1 1
1961/01/01 02:00:00           1.7800 1 1
1961/01/01 03:00:00           1.8500 1 1
1961/01/01 04:00:00           1.9000 1 1
1961/01/01 05:00:00           1.9100 1 1
1961/01/01 06:00:00           1.8300 1 1
1961/01/01 07:00:00           1.6800 1 1
1961/01/01 08:00:00           1.4900 1 1
1961/01/01 09:00:00           1.2800 1 1
1961/01/01 10:00:00           1.0700 1 1
1961/01/01 11:00:00           0.9100 1 1
1961/01/01 12:00:00           0.8300 1 1
1961/01/01 13:00:00           0.8100 1 1
1961/01/01 14:00:00           0.8800 1 1
```

2. Example of an extract of a UHSLC JASL hourly data file

```
806NOUAKCHO2007  LAT=18 06. N  LONG=015 57. W  TIMEZONE=GMT
806NOUAKCHO200701011 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999
806NOUAKCHO200701012 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999
806NOUAKCHO200701021 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999
806NOUAKCHO200701022 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999
806NOUAKCHO200701031 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999
806NOUAKCHO200701032 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999
806NOUAKCHO200701041 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999
806NOUAKCHO200701042 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999
806NOUAKCHO200701051 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999 9999
806NOUAKCHO200701052 9999 9999  791  579  315  290  375  568  837 1128 1369 1448
806NOUAKCHO200701061 1436 1314 1003  715  448  265  232  333  496  727  958 1121
806NOUAKCHO200701062 1219 1111  942  721  442  336  321  438  694  954 1233 1411
806NOUAKCHO200701071 1449 1435 1259  972  677  414  256  318  430  577  883 1076
806NOUAKCHO200701072 1196 1189 1062  896  634  417  384  398  585  818 1101 1271
806NOUAKCHO200701081 1419 1458 1390 1113  862  624  413  353  427  535  750  969
806NOUAKCHO200701082 1116 1198 1182 1025  835  639  485  448  549  746  971 1151
806NOUAKCHO200701091 1317 1398 1415 1267 1034  799  588  442  448  559  733  944
806NOUAKCHO200701092 1092 1164 1234 1146  980  828  650  523  569  654  829 1038
806NOUAKCHO200701101 1195 1337 1394 1277 1199  963  832  597  535  575  700  850
806NOUAKCHO200701102  997 1070 1229 1203 1103  994  869  630  601  619  760  909
806NOUAKCHO200701111 1071 1171 1285 1286 1209 1031  896  722  551  524  584  716
806NOUAKCHO200701112  819  925 1013 1137 1046 1024  863  780  688  636  677  774
806NOUAKCHO200701121  858 1067 1065 1158 1111 1070  928  834  679  572  529  628
806NOUAKCHO200701122  758  825  977 1053 1095 1095 1029  945  829  734  736  743
806NOUAKCHO200701131  822  901  997 1099 1155 1090 1032  921  795  746  607  609
```

```
806NOUAKCHO200701132  662  741  870  955 1031 1090 1105 1083  993  888  795  740
```

3. <u>Example NetCDF CDL from NOAA</u>

```
netcdf noaa_sanfrancisco_200502 {
dimensions:
        row = 108 ;
        stationID_strlen = 7 ;
        stationName_strlen = 13 ;
        state_strlen = 2 ;
        shefID_strlen = 5 ;
        deployment_strlen = 11 ;
        datum_strlen = 4 ;
        type_strlen = 1 ;
variables:
        char stationID(row, stationID_strlen) ;
                stationID:_Encoding = "ISO-8859-1" ;
                stationID:cf_role = "timeseries_id" ;
                stationID:comment = "Queries for data MUST include \"stationID=\"." ;
                stationID:ioos_category = "Identifier" ;
                stationID:long_name = "Station ID" ;
        char stationName(row, stationName_strlen) ;
                stationName:_Encoding = "ISO-8859-1" ;
                stationName:ioos_category = "Identifier" ;
                stationName:long_name = "Station Name" ;
        char state(row, state_strlen) ;
                state:_Encoding = "ISO-8859-1" ;
                state:ioos_category = "Location" ;
                state:long_name = "State" ;
        double dateEstablished(row) ;
                dateEstablished:actual_range = -3645043200., -3645043200. ;
                dateEstablished:ioos_category = "Time" ;
                dateEstablished:long_name = "Date Established" ;
                dateEstablished:time_origin = "01-JAN-1970 00:00:00" ;
                dateEstablished:time_precision = "1970-01-01" ;
                dateEstablished:units = "seconds since 1970-01-01T00:00:00Z" ;
        char shefID(row, shefID_strlen) ;
                shefID:_Encoding = "ISO-8859-1" ;
                shefID:comment = "A.K.A. NWS Location Identifier (NWSLI)" ;
                shefID:ioos_category = "Identifier" ;
                shefID:long_name = "SHEF ID" ;
        char deployment(row, deployment_strlen) ;
                deployment:_Encoding = "ISO-8859-1" ;
                deployment:ioos_category = "Identifier" ;
                deployment:long_name = "Deployment Designation" ;
        double longitude(row) ;
                longitude:_CoordinateAxisType = "Lon" ;
                longitude:actual_range = -122.4659, -122.4659 ;
```

```
                longitude:axis = "X" ;
                longitude:ioos_category = "Location" ;
                longitude:long_name = "Longitude" ;
                longitude:standard_name = "longitude" ;
                longitude:units = "degrees_east" ;
        double latitude(row) ;
                latitude:_CoordinateAxisType = "Lat" ;
                latitude:actual_range = 37.8063, 37.8063 ;
                latitude:axis = "Y" ;
                latitude:ioos_category = "Location" ;
                latitude:long_name = "Latitude" ;
                latitude:standard_name = "latitude" ;
                latitude:units = "degrees_north" ;
        double time(row) ;
                time:_CoordinateAxisType = "Time" ;
                time:actual_range = 1107233400., 1109628960. ;
                time:axis = "T" ;
                time:comment = "Queries for data MUST include \"time>=\" and \"time<=\"." ;
                time:ioos_category = "Time" ;
                time:long_name = "Time" ;
                time:standard_name = "time" ;
                time:time_origin = "01-JAN-1970 00:00:00" ;
                time:units = "seconds since 1970-01-01T00:00:00Z" ;
        char datum(row, datum_strlen) ;
                datum:_Encoding = "ISO-8859-1" ;
                datum:comment = "This dataset only returns data with the MLLW (Mean Lower-Low Water) datum." ;
                datum:ioos_category = "Location" ;
                datum:long_name = "Datum" ;
        float waterLevel(row) ;
                waterLevel:actual_range = -0.464f, 2.119f ;
                waterLevel:ioos_category = "Sea Level" ;
                waterLevel:long_name = "Water Level" ;
                waterLevel:standard_name = "sea_surface_height_amplitude_due_to_geocentric_ocean_tide" ;
                waterLevel:units = "m" ;
        char type(row, type_strlen) ;
                type:_Encoding = "ISO-8859-1" ;
                type:comment = "H=High water, L=Low water" ;
                type:ioos_category = "Sea Level" ;
                type:long_name = "Designation of Water Level Height" ;

// global attributes:
                :cdm_data_type = "TimeSeries" ;
                :cdm_timeseries_variables = "stationID, stationName, state, dateEstablished, shefID, deployment, longitude, latitude, datum" ;
                :Conventions = "COARDS, CF-1.6, ACDD-1.3" ;
                :creator_email = "COOPS.IOOS@noaa.gov" ;
                :creator_name = "NOAA NOS COOPS" ;
                :creator_type = "institution" ;
                :creator_url = "http://tidesandcurrents.noaa.gov/" ;
                :defaultDataQuery = "&time%3E=now&time%3C=now+7days&stationName=%22San%20Francisco%22&datum=%22MLLW%22" ;
```

```
                :defaultGraphQuery =
"time,waterLevel&time%3E=now&time%3C=now+7days&stationName=%22San%20Francisco%22&datum=%22MLLW%22&.draw=lines" ;
                :Easternmost_Easting = -122.4659 ;
                :featureType = "TimeSeries" ;
                :geospatial_lat_max = 37.8063 ;
                :geospatial_lat_min = 37.8063 ;
                :geospatial_lat_units = "degrees_north" ;
                :geospatial_lon_max = -122.4659 ;
                :geospatial_lon_min = -122.4659 ;
                :geospatial_lon_units = "degrees_east" ;
                :history = "2020-03-24T15:31:39Z https://opendap.co-
ops.nos.noaa.gov/axis/webservices/highlowtidepred/plain/response.jsp?unit=1&timeZone=1&datum=0&metadata=yes&Submit=Submit\n",
                        "2020-03-24T15:31:39Z
http://coastwatch.pfeg.noaa.gov/erddap/tabledap/nosCoopsWLTPHL.nc?stationID%2CstationName%2Cstate%2CdateEstablished%2CshefID%2Cdeployment%2Clongi
tude%2Clatitude%2Ctime%2Cdatum%2CwaterLevel%2Ctype&stationName=%22San%20Francisco%22&time%3E=2005-02-01T00%3A00%3A00Z&time%3C=2005-02-
28T23%3A59%3A59Z&datum=%22MLLW%22" ;
                :id = "nosCoopsWLTPHL" ;
                :infoUrl = "http://tidesandcurrents.noaa.gov/" ;
                :institution = "NOAA NOS CO-OPS" ;
                :keywords = "amplitude, co-ops, coastal, data, datum, deployment, designation, due, Earth Science > Oceans > Coastal Processes >
Tidal Height, Earth Science > Oceans > Sea Surface Topography > Sea Surface Height, Earth Science > Oceans > Tides > Tidal Height, geocentric,
height, high, identifier, level, low, name, noaa, nos, ocean, oceans, ops, prediction, processes, sea, sea level,
sea_surface_height_amplitude_due_to_geocentric_ocean_tide, seawater, shef_id, shefID, station, surface, tidal, tide, tides, time, topography,
water" ;
                :keywords_vocabulary = "GCMD Science Keywords" ;
                :license = "The official Tide and Tidal Current prediction tables are published annually on\n",
                        "October 1, for the following calendar year. Tide and Tidal Current predictions\n",
                        "generated prior to the publishing date of the official tables are subject to\n",
                        "change. The enclosed data are based upon the latest information available as of\n",
                        "the date of your request. Tide and Tidal Current predictions generated may\n",
                        "differ from the official predictions if information for the station requested\n",
                        "has been updated since the publishing date of the official tables.\n",
                        "\n",
                        "The data may be used and redistributed for free but is not intended\n",
                        "for legal use, since it may contain inaccuracies. Neither the data\n",
                        "Contributor, ERD, NOAA, nor the United States Government, nor any\n",
                        "of their employees or contractors, makes any warranty, express or\n",
                        "implied, including warranties of merchantability and fitness for a\n",
                        "particular purpose, or assumes any legal liability for the accuracy,\n",
                        "completeness, or usefulness, of this information." ;
                :Northernmost_Northing = 37.8063 ;
                :sourceUrl = "https://opendap.co-
ops.nos.noaa.gov/axis/webservices/highlowtidepred/plain/response.jsp?unit=1&timeZone=1&datum=0&metadata=yes&Submit=Submit" ;
                :Southernmost_Northing = 37.8063 ;
                :standard_name_vocabulary = "CF Standard Name Table v70" ;
                :subsetVariables = "stationID, stationName, state, dateEstablished, shefID, deployment, longitude, latitude, datum" ;
                :summary = "This dataset has High Low Tide Predictions from NOAA NOS Center for Operational\n",
                        "Oceanographic Products and Services (CO-OPS).\n",
                        "\n",
```

```
                      "The official Tide and Tidal Current prediction tables are published annually on\n",
                      "October 1, for the following calendar year. Tide and Tidal Current predictions\n",
                      "generated prior to the publishing date of the official tables are subject to\n",
                      "change. The enclosed data are based upon the latest information available as of\n",
                      "the date of your request. Tide and Tidal Current predictions generated may\n",
                      "differ from the official predictions if information for the station requested\n",
                      "has been updated since the publishing date of the official tables.\n",
                      "\n",
                      "WARNING:\n",
                      "* Queries for data MUST include stationID=, time>= and time<=.\n",
                      "  Queries MUST be for less than 30 days worth of data.\n",
                      "* This dataset only returns data for the MLLW (Mean Lower-Low Water) datum.\n",
                      "* The data source isn\'t completely reliable. If your request returns no data\n",
                      "  when you think it should:\n",
                      "  * Try revising the request (e.g., a different time range).\n",
                      "  * The list of stations offering this data may be incorrect.\n",
                      "  * Sometimes a station or the entire data service is unavailable.\n",
                      "      Wait a while and try again." ;
              :time_coverage_end = "2005-02-28T22:16:00Z" ;
              :time_coverage_start = "2005-02-01T04:50:00Z" ;
              :title = "NOS CO-OPS Water Level Data, High Low Tide Prediction" ;
              :Westernmost_Easting = -122.4659 ;
}
```

4. Example NetCDF CDL from UHSLC

```
netcdf uhslc_sanfrancisco_2005 {
dimensions:
        row = 8760 ;
        station_name_strlen = 17 ;
        station_country_strlen = 30 ;
        ssc_id_strlen = 4 ;
variables:
        short sea_level(row) ;
                sea_level:_FillValue = -32767s ;
                sea_level:actual_range = 1191s, 4298s ;
                sea_level:long_name = "relative sea level" ;
                sea_level:platform = "station_name, station_country, station_country_code, uhslc_id, gloss_id, ssc_id" ;
                sea_level:source = "in situ tide gauge water level observations" ;
                sea_level:units = "millimeters" ;
        double time(row) ;
                time:_CoordinateAxisType = "Time" ;
                time:actual_range = 1104537600., 1136069999.971 ;
                time:axis = "T" ;
                time:ioos_category = "Time" ;
                time:long_name = "Time" ;
                time:standard_name = "time" ;
                time:time_origin = "01-JAN-1970 00:00:00" ;
                time:units = "seconds since 1970-01-01T00:00:00Z" ;
        float latitude(row) ;
```

```
		latitude:_CoordinateAxisType = "Lat" ;
		latitude:actual_range = 37.807f, 37.807f ;
		latitude:axis = "Y" ;
		latitude:colorBarMaximum = 90. ;
		latitude:colorBarMinimum = -90. ;
		latitude:ioos_category = "Location" ;
		latitude:long_name = "Latitude" ;
		latitude:standard_name = "latitude" ;
		latitude:units = "degrees_north" ;
		latitude:valid_max = 90.f ;
		latitude:valid_min = -90.f ;
	float longitude(row) ;
		longitude:_CoordinateAxisType = "Lon" ;
		longitude:actual_range = 237.535f, 237.535f ;
		longitude:axis = "X" ;
		longitude:colorBarMaximum = 180. ;
		longitude:colorBarMinimum = -180. ;
		longitude:ioos_category = "Location" ;
		longitude:long_name = "Longitude" ;
		longitude:standard_name = "longitude" ;
		longitude:units = "degrees_east" ;
		longitude:valid_max = 360.f ;
		longitude:valid_min = 0.f ;
	char station_name(row, station_name_strlen) ;
		station_name:_Encoding = "ISO-8859-1" ;
		station_name:long_name = "station name" ;
	char station_country(row, station_country_strlen) ;
		station_country:_Encoding = "ISO-8859-1" ;
		station_country:long_name = "station country (ISO 3166-1)" ;
	short station_country_code(row) ;
		station_country_code:_FillValue = 0s ;
		station_country_code:actual_range = 840s, 840s ;
		station_country_code:comment = "These are 3-digit country codes (e.g., 003) stored as integers." ;
		station_country_code:long_name = "station country code (ISO 3166-1 numeric)" ;
	short record_id(row) ;
		record_id:actual_range = 5511s, 5511s ;
		record_id:long_name = "unique identifier for each record (i.e., station and version) in the database" ;
	short uhslc_id(row) ;
		uhslc_id:actual_range = 551s, 551s ;
		uhslc_id:cf_role = "timeseries_id" ;
		uhslc_id:colorBarMaximum = 100. ;
		uhslc_id:colorBarMinimum = 0. ;
		uhslc_id:long_name = "unique station ID number used by the University of Hawaii Sea Level Center (UHSLC)" ;
	char version(row) ;
		version:actual_range = "AA" ;
		version:comment = "The station version is a letter from A to Z differentiating segments of a station record that cannot be linked
to a common benchmark." ;
		version:long_name = "station version" ;
	short gloss_id(row) ;
```

```
                gloss_id:_FillValue = 0s ;
                gloss_id:actual_range = 158s, 158s ;
                gloss_id:colorBarMaximum = 100. ;
                gloss_id:colorBarMinimum = 0. ;
                gloss_id:long_name = "unique station ID number used by the WMO/IOC Global Sea Level Observing System (GLOSS)" ;
        char ssc_id(row, ssc_id_strlen) ;
                ssc_id:_Encoding = "ISO-8859-1" ;
                ssc_id:comment = "Note that SSC IDs vary in length. The IDs are padded with space characters to produce the retangular character
matrix provided here." ;
                ssc_id:long_name = "unique station ID code in the Sealevel Station Catalog (SSC) produced by the WMO/IOC Sea Level Monitoring
Facility (VLIZ)" ;
        short decimation_method(row) ;
                decimation_method:actual_range = 3s, 3s ;
                decimation_method:colorBarMaximum = 5. ;
                decimation_method:colorBarMinimum = 0. ;
                decimation_method:flag_meanings = "filtered, average, spot readings, other" ;
                decimation_method:flag_values = "1, 2, 3, 4" ;
                decimation_method:long_name = "decimation method" ;
        char reference_code(row) ;
                reference_code:actual_range = "RR" ;
                reference_code:colorBarMaximum = 5. ;
                reference_code:colorBarMinimum = 0. ;
                reference_code:flag_meanings = "data referenced to datum, data not referenced to datum" ;
                reference_code:flag_values = "R, X" ;
                reference_code:long_name = "reference code" ;
        short reference_offset(row) ;
                reference_offset:actual_range = 0s, 0s ;
                reference_offset:comment = "This is a constant offset to be added to each data value to make the data relative to the tide staff
zero or primary datum." ;
                reference_offset:long_name = "reference offset" ;
                reference_offset:units = "millimeters" ;

// global attributes:
                :acknowledgement = "The JASL/UHSLC Research Quality Data Set is supported by the National Oceanic and Atmospheric Administration
(NOAA) via the National Centers for Environmental Information (NCEI) and the Office of Climate Observations (OCO)." ;
                :acknowledgment = "The JASL/UHSLC Research Quality Data Set is supported by the National Oceanic and Atmospheric Administration
(NOAA) via the National Centers for Environmental Information (NCEI) and the Office of Climate Observations (OCO)." ;
                :cdm_data_type = "TimeSeries" ;
                :cdm_timeseries_variables = "uhslc_id, latitude, longitude" ;
                :Conventions = "CF-1.6, ACDD-1.3, COARDS" ;
                :date_created = "2019-07-12T23:52:10Z" ;
                :Easternmost_Easting = 237.535f ;
                :featureType = "TimeSeries" ;
                :geospatial_lat_max = 37.807f ;
                :geospatial_lat_min = 37.807f ;
                :geospatial_lat_units = "degrees_north" ;
                :geospatial_lon_max = 237.535f ;
                :geospatial_lon_min = 237.535f ;
                :geospatial_lon_units = "degrees_east" ;
```

```
		:history = "2020-03-24T15:23:10Z (local files)\n",
			"2020-03-24T15:23:10Z
https://uhslc.soest.hawaii.edu/erddap/tabledap/global_hourly_rqds.nc?sea_level%2Ctime%2Clatitude%2Clongitude%2Cstation_name%2Cstation_country%2Cs
tation_country_code%2Crecord_id%2Cuhslc_id%2Cversion%2Cgloss_id%2Cssc_id%2Cdecimation_method%2Creference_code%2Creference_offset&time%3E=2005-01-
01T00%3A00%3A00Z&time%3C=2005-12-31T23%3A59%3A59Z&station_name=%22San%20Francisco%2C%20CA%22" ;
		:id = "global_hourly_rqds" ;
		:infoUrl = "https://uhslc.soest.hawaii.edu/data/" ;
		:institution = "University of Hawaii Sea Level Center" ;
		:keywords = "3166-1, archive, center, centers, code, commission, control, country, data, database, decimation, decimation_method,
each, environmental, facility, gauge, global, gloss, gloss_id, hawaii, hour, hourly, i.e., identifier, information, intergovernmental, ioc, iso,
jasl, jasl/uhslc, joint, latitude, level, longitude, meteorological, method, monitoring, name, national, ncei, nesdis, noaa, number, numeric,
observing, oceanographic, offset, organisation, produced, quality, record, record_id, reference, reference_code, reference_offset, relative,
research, rqds, sea, sea_level, sealevel, set, ssc, ssc_id, station, station_country, station_country_code, station_name, system, tide, time,
uhslc, uhslc_id, unique, university, used, version, vliz, wmo, wmo/ioc, world" ;
		:license = "The data may be used and redistributed for free but is not intended\n",
			"for legal use, since it may contain inaccuracies. Neither the data\n",
			"Contributor, ERD, NOAA, nor the United States Government, nor any\n",
			"of their employees or contractors, makes any warranty, express or\n",
			"implied, including warranties of merchantability and fitness for a\n",
			"particular purpose, or assumes any legal liability for the accuracy,\n",
			"completeness, or usefulness, of this information." ;
		:Northernmost_Northing = 37.807f ;
		:processing_level = "The JASL receives hourly data from regional and national sea level networks operating world-wide. JASL RQDS
data undergo a level 1 quality assessment focused on variability (e.g., unit and timing evaluation, outlier detection, combination of multiple
channels into a primary channel, etc.) followed by a level 2 quality assessment focused on datum stability (e.g., tide gauge datum evaluation,
assessment of level ties to tide gauge benchmarks, comparison with nearby stations, etc.)." ;
		:publisher_email = "philiprt@hawaii.edu, markm@soest.hawaii.edu" ;
		:publisher_name = "University of Hawaii Sea Level Center (UHSLC)" ;
		:publisher_type = "group" ;
		:publisher_url = "http://uhslc.soest.hawaii.edu" ;
		:sourceUrl = "(local files)" ;
		:Southernmost_Northing = 37.807f ;
		:standard_name_vocabulary = "CF Standard Name Table v55" ;
		:subsetVariables = "latitude, longitude, station_name, station_country, station_country_code, record_id, uhslc_id, version,
gloss_id, ssc_id, decimation_method, reference_code, reference_offset" ;
		:summary = "The Joint Archive for Sea Level (JASL) Research Quality Data Set (RQDS) is a collaboration between the University of
Hawaii Sea Level Center (UHSLC) and the World Data Center for Oceanography of the National Centers for Environmental Information (NCEI), National
Oceanic and Atmospheric Administration (NOAA). The objective of the JASL RQDS is to assemble a well-documented, quality-controlled archive of
hourly and daily sea level values that is appropriate for scientific research applications. The JASL RQDS is the largest global collection of
quality-controlled hourly sea level data, and ongoing efforts seek to acquire new sites and uncover historic records as available." ;
		:time_coverage_end = "2005-12-31T22:59:59Z" ;
		:time_coverage_start = "2005-01-01T00:00:00Z" ;
		:title = "JASL/UHSLC Research Quality Tide Gauge Data (hourly)" ;
		:Westernmost_Easting = 237.535f ;
}
```

5. Example NetCDF CDL from CMEMS

```
netcdf cmems_newlyn {
dimensions:
        TIME = 250068 ;
        DEPTH = 1 ;
        LATITUDE = 250068 ;
        LONGITUDE = 250068 ;
        POSITION = 250068 ;
variables:
        double TIME(TIME) ;
                TIME:long_name = "Time"
                TIME:standard_name = "time" ;
                TIME:units = "days since 1950-01-01T00:00:00Z" ;
                TIME:_FillValue = 9.96920996838687e+36 ;
                TIME:valid_min = -90000.f ;
                TIME:valid_max = 90000.f ;
                TIME:QC_indicator = 1b ;
                TIME:QC_procedure = 1b ;
                TIME:uncertainty = " " ;
                TIME:comment = " " ;
                TIME:axis = "T" ;
        byte TIME_QC(TIME) ;
                TIME_QC:long_name = "quality flag" ;
                TIME_QC:conventions = "OceanSITES reference table 2" ;
                TIME_QC:_FillValue = -127b ;
                TIME_QC:valid_min = 0b ;
                TIME_QC:valid_max = 9b ;
                TIME_QC:flag_values = 0b, 1b, 2b, 3b, 4b, 5b, 6b, 7b, 8b, 9b ;
                TIME_QC:flag_meanings = "no_qc_performed good_data probably_good_data bad_data_that_are_potentially_correctable bad_data
value_changed not_used nominal_value interpolated_value missing_value" ;
        float LATITUDE(LATITUDE) ;
                LATITUDE:long_name = "Latitude of each location" ;
                LATITUDE:standard_name = "latitude" ;
                LATITUDE:units = "degree_north" ;
                LATITUDE:_FillValue = 9.96921e+36f ;
                LATITUDE:valid_min = -90.f ;
                LATITUDE:valid_max = 90.f ;
                LATITUDE:QC_indicator = 1b ;
                LATITUDE:QC_procedure = 1b ;
                LATITUDE:uncertainty = " " ;
                LATITUDE:comment = " " ;
                LATITUDE:axis = "Y" ;
        float LONGITUDE(LONGITUDE) ;
                LONGITUDE:long_name = "Longitude of each location" ;
                LONGITUDE:standard_name = "longitude" ;
```

```
                LONGITUDE:units = "degree_east" ;
                LONGITUDE:_FillValue = 9.96921e+36f ;
                LONGITUDE:valid_min = -180.f ;
                LONGITUDE:valid_max = 180.f ;
                LONGITUDE:QC_indicator = 1b ;
                LONGITUDE:QC_procedure = 1b ;
                LONGITUDE:uncertainty = " " ;
                LONGITUDE:comment = " " ;
                LONGITUDE:axis = "X" ;
        byte POSITION_QC(POSITION) ;
                POSITION_QC:long_name = "quality flag" ;
                POSITION_QC:conventions = "OceanSITES reference table 2" ;
                POSITION_QC:_FillValue = -127b ;
                POSITION_QC:valid_min = 0b ;
                POSITION_QC:valid_max = 9b ;
                POSITION_QC:flag_values = 0b, 1b, 2b, 3b, 4b, 5b, 6b, 7b, 8b, 9b ;
                POSITION_QC:flag_meanings = "no_qc_performed good_data probably_good_data bad_data_that_are_potentially_correctable bad_data
value_changed not_used nominal_value interpolated_value missing_value" ;
        float DEPH(TIME, DEPTH) ;
                DEPH:long_name = "Depth" ;
                DEPH:standard_name = "depth" ;
                DEPH:units = "m" ;
                DEPH:positive = "down" ;
                DEPH:_FillValue = 9.96921e+36f ;
                DEPH:valid_min = -12000.f ;
                DEPH:valid_max = 12000.f ;
                DEPH:QC_indicator = 1b ;
                DEPH:QC_procedure = 1b ;
                DEPH:uncertainty = " " ;
                DEPH:comment = " " ;
                DEPH:axis = "Z" ;
                DEPH:reference = "sea_level" ;
        byte DEPH_QC(TIME, DEPTH) ;
                DEPH_QC:long_name = "quality flag" ;
                DEPH_QC:conventions = "OceanSITES reference table 2" ;
                DEPH_QC:_FillValue = -127b ;
                DEPH_QC:valid_min = 0b ;
                DEPH_QC:valid_max = 9b ;
                DEPH_QC:flag_values = 0b, 1b, 2b, 3b, 4b, 5b, 6b, 7b, 8b, 9b ;
                DEPH_QC:flag_meanings = "no_qc_performed good_data probably_good_data bad_data_that_are_potentially_correctable bad_data
value_changed not_used nominal_value interpolated_value missing_value" ;
        char DEPH_DM(TIME, DEPTH) ;
                DEPH_DM:long_name = "method of data processing" ;
                DEPH_DM:conventions = "OceanSITES reference table 5" ;
                DEPH_DM:flag_values = "R, P, D, M" ;
```

```
                DEPH_DM:flag_meanings = "real-time provisional delayed-mode mixed" ;
                DEPH_DM:_FillValue = " " ;

        int SLEV(TIME, DEPTH) ;
                SLEV:long_name = "Water surface height above a specific datum" ;
                SLEV:standard_name = "water_surface_height_above_reference_datum" ;
                SLEV:units = "m" ;
                SLEV:_FillValue = -2147483647 ;
                SLEV:QC_procedure = 1b ;
                SLEV:valid_min = -10000 ;
                SLEV:valid_max = 20000 ;
                SLEV:comment = " " ;
                SLEV:sensor_depth = 0.f ;
                SLEV:ancillary_variables = "SLEV_QC" ;
                SLEV:sensor_mount = " " ;
                SLEV:sensor_orientation = " " ;
                SLEV:DM_indicator = "R" ;
                SLEV:scale_factor = 0.001f ;
                SLEV:add_offset = 0.f ;
                SLEV:time_sampling = 60.f ;
                SLEV:sea_level_datum = "Admiralty Chart Datum" ;
                SLEV:processing_method = "filtered values" ;
        byte SLEV_QC(TIME, DEPTH) ;
                SLEV_QC:long_name = "quality flag" ;
                SLEV_QC:conventions = "OceanSITES reference table 2" ;
                SLEV_QC:_FillValue = -127b ;
                SLEV_QC:valid_min = 0b ;
                SLEV_QC:valid_max = 9b ;
                SLEV_QC:flag_values = 0b, 1b, 2b, 3b, 4b, 5b, 6b, 7b, 8b, 9b ;
                SLEV_QC:flag_meanings = "no_qc_performed good_data probably_good_data bad_data_that_are_potentially_correctable bad_data
value_changed not_used nominal_value interpolated_value missing_value" ;
        char SLEV_DM(TIME, DEPTH) ;
                SLEV_DM:long_name = "method of data processing" ;
                SLEV_DM:conventions = "OceanSITES reference table 5" ;
                SLEV_DM:flag_values = "R, P, D, M" ;
                SLEV_DM:flag_meanings = "real-time provisional delayed-mode mixed" ;
                SLEV_DM:_FillValue = " " ;

// global attributes:
                :data_type = "OceanSITES time-series data" ;
                :format_version = "1.2" ;
                :platform_code = "Newlyn" ;
                :date_update = "2020-03-03T10:36:03Z" ;
                :institution = "National Oceanography Centre (United_Kingdom)" ;
                :institution_edmo_code = "2424" ;
                :site_code = " " ;
```

```
                :wmo_platform_code = " " ;
                :source = "mooring" ;
                :source_platform_category_code = "48" ;
                :history = "2020-03-03T10:36:03Z: Creation" ;
                :data_mode = "R" ;
                :quality_control_indicator = "6" ;
                :quality_index = "A" ;
                :references = "http://www.oceansites.org, http://marine.copernicus.eu, http://noc.ac.uk/, http://www.puertos.es" ;
                :comment = " " ;
                :Conventions = "CF-1.6 OceanSITES-Manual-1.2 Copernicus-InSituTAC-SRD-1.4 Copernicus-InSituTAC-ParametersList-3.1.0" ;
                :netcdf_version = "netCDF-4 classic model" ;
                :title = "IBI - NRT in situ Observations" ;
                :summary = "" ;
                :naming_authority = "OceanSITES" ;
                :id = "IR_TS_TG_Newlyn" ;
                :cdm_data_type = "Time-series" ;
                :area = "North Atlantic Ocean" ;
                :geospatial_lat_min = "50.1" ;
                :geospatial_lat_max = "50.1" ;
                :geospatial_lon_min = "-5.53333" ;
                :geospatial_lon_max = "-5.53333" ;
                :geospatial_vertical_min = " " ;
                :geospatial_vertical_max = " " ;
                :time_coverage_start = "1990-01-01T02:00:00Z" ;
                :time_coverage_end = "2020-02-29T23:45:00Z" ;
                :institution_references = "http://noc.ac.uk/" ;
                :contact = "ljr@bodc.ac.uk,mar@puertos.es" ;
                :author = "Marta de Alfonso" ;
                :data_assembly_center = "Puertos del Estado" ;
                :pi_name = "PdE" ;
                :distribution_statement = "These data follow Copernicus standards; they are public and free of charge. User assumes all risk for
use of data. User must display citation in any publication or product using data. User must contact PI prior to any commercial use of data." ;
                :citation = "These data were collected and made freely available by the Copernicus project and the programs that contribute to it"
;
                :update_interval = "yearly" ;
                :qc_manual = "OceanSITES User\'s Manual v1.2" ;
                :last_latitude_observation = 50.1 ;
                :last_longitude_observation = -5.5333 ;
                :last_date_observation = "2020-02-29T23:00:00Z" ;


}
```

ANNEX III

ACRONYMS

ACDD	Attribute Convention for Data Discovery
Argo	A global array of profiling floats
API	Application Programming Interface
ARP	Antenna Reference Point
ASCII	American Standard Code for Information Interchange
AtlantOS	Optimising and Enhancing the Integrated Atlantic Ocean Observing Systems – EU funded project
BODC	British Oceanographic Data Centre (UK)
BOM	Bureau of Meteorology (Australia)
CARIBE-EWS	Tsunami and Other Coastal Hazards Warning System for the Caribbean and Adjacent Regions
CDL	Common Data Language
CF	Climate and Forecast
CGNSS	Continuous Global Navigation Satellite System
CHS	Canadian Hydrographic Service
CMEMS	Copernicus Marine Environment Monitoring Service
CNRS/INSU	Centre national de la recherche scientifique/National Institute for Earth Sciences and Astronomy (France)
DAC	Dynamic Atmospheric Correction
DFO	Department of Fisheries and Oceans (Canada)
DMPA	Data Management Programme Area (a component of JCOMM)
DOI	Digital Object Identifiers
EDMED	European Directory of Marine Environmental Data Sets
EDMO	European Directory of Marine Organisations
EMODnet	European Marine Observation and Data Network
ENSO	El Niño–Southern Oscillation
EOF	Empirical Orthogonal Function
ESEAS-RI	European Sea Level Service Research Infrastructure
EuroGOOS	European Global Ocean Observing System
FAIR	Findable, Accessible, Interoperable, and Re-useable
FTP	File Transfer Protocol
GCMD	Global Change Master Directory
GCN	Global Core Network
GEOSS	Global Earth Observation System of Systems
GESLA	Global Extreme Sea Level Analysis
GFZ	GeoForschungsZentrum (German Research Centre for Geosciences)

GLOSS	Global Sea Level Observing System – 'Global Level of the Sea Surface'
GNSS	Global Navigation Satellite System
GOOS	Global Ocean Observing System
GTS	Global Telecommunication System (global network for the transmission of meteorological data)
GTSPP	Global Temperature and Salinity Profile Project
IOS	Institute of Ocean Sciences (Canada)
ICES DIG	International Council for the Exploration of the Sea's Data and Information Group
IGS	International GNSS (Global Navigation Satellite System) Service
IHO	International Hydrographic Organization
IOC	Intergovernmental Oceanographic Commission (of UNESCO)
IODE	International Oceanographic Data and Information Exchange
IOOS	US Integrated Ocean Observing System
IOTWMS	Indian Ocean Tsunami Warning and Mitigation System
ISO	International Standards Organization
ISPRA	Istituto Superiore per la Protezione e la Ricerca Ambientale (Italy)
JASL	Joint Archive for Sea Level
JCOMM	Joint WMO-IOC Technical Commission for Oceanography and Marine Meteorology
KHOA	Korea Hydrographic and Oceanographic Agency (Republic of Korea)
LTT	Long Term Trend
MHW	Mean High Waters
MLW	Mean Low Waters
MTL	Mean Tide Level
MSL	Mean Sea Level
MyOcean	EU funded projects for "Development and pre-operational validation of upgraded GMES Marine Core Services and capabilities"
NAO	North Atlantic Oscillation
NEAMTWS	North-East Atlantic, Mediterranean and adjacent seas Tsunami Warning System
NetCDF	Network Common Data Form
NIWA	National Institute of Water and Atmospheric Research (New Zealand)
NOC	National Oceanography Centre (UK)
NOAA	National Oceanic and Atmospheric Administration (USA)
NRT	Near Real Time (data)
NTR	Non-Tidal Residual
PdE	Puertos del Estado (Spain)
PSMSL	Permanent Service for Mean Sea Level

PTWS	Pacific Tsunami Warning System
QARTOD	Quality Assurance / Quality Control of Real Time Oceanographic Data
QC	Quality Control
REDMAR	Spanish Harbours Authority Sea Level Network
SANHO	South African Navy Hydrographic Office
SeaDataNet	Pan-European infrastructure for ocean and marine data management
SLSMF	Sea Level Monitoring Facility
SELENE	SEa LEvel NEar-real time quality control and processing
SNHT	Standard Normal Homogeneity Test
SONEL	Système d'Observation du Niveau des Eaux Littorales (at University of La Rochelle) (France)
SMA	Swedish Maritime Administration (Sweden)
TASK	Tidal Analysis Software Kit (of NOC)
TGBM	Tide Gauge Benchmark
TIGA	TIde GAuge (project of the IGS), a Tide Gauge Benchmark Monitoring Project
TIRA	Tidal Institute Recursive Analysis
TOGA	Tropical Ocean Global Atmosphere project
T_Tide	T_Tide = Classical tidal harmonic analysis software
UTIDE	UTIDE = Unified Tidal Analysis and Prediction software
UHSLC	University of Hawaii Sea Level Center (USA)
UNAM	Universidad Nacional Autónoma de México (Mexico)
UNESCO	United Nations Educational, Scientific and Cultural Organization
UTC	Coordinated Universal Time
VLIZ	Flanders Marine Institute (Belgium)
WMO	World Meteorological Organization
WOCE	World Ocean Circulation Experiment

IOC Manuals and Guides

No.	Title
1 rev. 2	Guide to IGOSS Data Archives and Exchange (BATHY and TESAC). 1993. 27 pp. (English, French, Spanish, Russian)
2	International Catalogue of Ocean Data Station. 1976. *(Out of stock)*
3 rev. 3	Guide to Operational Procedures for the Collection and Exchange of JCOMM Oceanographic Data. Third Revised Edition, 1999. 38 pp. (English, French, Spanish, Russian)
4	Guide to Oceanographic and Marine Meteorological Instruments and Observing Practices. 1975. 54 pp. (English)
5 rev. 2	Guide for Establishing a National Oceanographic Data Centre. Second Revised Edition, 2008. 27 pp. (English) *(Electronic only)*
6 rev.	Wave Reporting Procedures for Tide Observers in the Tsunami Warning System. 1968. 30 pp. (English)
7	Guide to Operational Procedures for the IGOSS Pilot Project on Marine Pollution (Petroleum) Monitoring. 1976. 50 pp. (French, Spanish)
8	*(Superseded by IOC Manuals and Guides No. 16)*
9 rev.	Manual on International Oceanographic Data Exchange. (Fifth Edition). 1991. 82 pp. (French, Spanish, Russian)
9 Annex I	*(Superseded by IOC Manuals and Guides No. 17)*
9 Annex II	Guide for Responsible National Oceanographic Data Centres. 1982. 29 pp. (English, French, Spanish, Russian)
10	*(Superseded by IOC Manuals and Guides No. 16)*
11	The Determination of Petroleum Hydrocarbons in Sediments. 1982. 38 pp. (French, Spanish, Russian)
12	Chemical Methods for Use in Marine Environment Monitoring. 1983. 53 pp. (English)
13	Manual for Monitoring Oil and Dissolved/Dispersed Petroleum Hydrocarbons in Marine Waters and on Beaches. 1984. 35 pp. (English, French, Spanish, Russian)
14	Manual on Sea-Level Measurements and Interpretation. (English, French, Spanish, Russian)
	Vol. I: Basic Procedure. 1985. 83 pp. (English)
	Vol. II: Emerging Technologies. 1994. 72 pp. (English)
	Vol. III: Reappraisals and Recommendations as of the year 2000. 2002. 55 pp. (English)
	Vol. IV: An Update to 2006. 2006. 78 pp. (English, Arab)
	Vol. V: Radar Gauges. 2016. 100 pp. and Supplement: Practical Experiences. 100 pp. (English, French, Russian, Spanish)
15	Operational Procedures for Sampling the Sea-Surface Microlayer. 1985. 15 pp. (English)
16	Marine Environmental Data Information Referral Catalogue. Third Edition. 1993. 157 pp. (Composite English/French/Spanish/Russian)
17	GF3: A General Formatting System for Geo-referenced Data
	Vol. 1: Introductory Guide to the GF3 Formatting System. 1993. 35 pp. (English, French, Spanish, Russian)
	Vol. 2: Technical Description of the GF3 Format and Code Tables. 1987. 111 pp. (English, French, Spanish, Russian)
	Vol. 3: Standard Subsets of GF3. 1996. 67 pp. (English)
	Vol. 4: User Guide to the GF3-Proc Software. 1989. 23 pp. (English, French, Spanish, Russian)

No.	Title
	Vol. 5: Reference Manual for the GF3-Proc Software. 1992. 67 pp. (English, French, Spanish, Russian)
	Vol. 6: Quick Reference Sheets for GF3 and GF3-Proc. 1989. 22 pp. (English, French, Spanish, Russian)
18	User Guide for the Exchange of Measured Wave Data. 1987. 81 pp. (English, French, Spanish, Russian)
19	Guide to IGOSS Specialized Oceanographic Centres (SOCs). 1988. 17 pp. (English, French, Spanish, Russian)
20	Guide to Drifting Data Buoys. 1988. 71 pp. (English, French, Spanish, Russian)
21	*(Superseded by IOC Manuals and Guides No. 25)*
22 rev.	GTSPP Real-time Quality Control Manual, First revised edition. 2010. 145 pp. (English)
23	Marine Information Centre Development: An Introductory Manual. 1991. 32 pp. (English, French, Spanish, Russian)
24	Guide to Satellite Remote Sensing of the Marine Environment. 1992. 178 pp. (English)
25	Standard and Reference Materials for Marine Science. Revised Edition. 1993. 577 pp. (English)
26	Manual of Quality Control Procedures for Validation of Oceanographic Data. 1993. 436 pp. (English)
27	Chlorinated Biphenyls in Open Ocean Waters: Sampling, Extraction, Clean-up and Instrumental Determination. 1993. 36 pp. (English)
28	Nutrient Analysis in Tropical Marine Waters. 1993. 24 pp. (English)
29	Protocols for the Joint Global Ocean Flux Study (JGOFS) Core Measurements. 1994. 178 pp . (English)
30	MIM Publication Series:
	Vol. 1: Report on Diagnostic Procedures and a Definition of Minimum Requirements for Providing Information Services on a National and/or Regional Level. 1994. 6 pp. (English)
	Vol. 2: Information Networking: The Development of National or Regional Scientific Information Exchange. 1994. 22 pp. (English)
	Vol. 3: Standard Directory Record Structure for Organizations, Individuals and their Research Interests. 1994. 33 pp. (English)
31	HAB Publication Series:
	Vol. 1: Amnesic Shellfish Poisoning. 1995. 18 pp. (English)
32	Oceanographic Survey Techniques and Living Resources Assessment Methods. 1996. 34 pp. (English)
33	Manual on Harmful Marine Microalgae. 1995. (English) [superseded by a sale publication in 2003, 92-3-103871-0. UNESCO Publishing]
34	Environmental Design and Analysis in Marine Environmental Sampling. 1996. 86 pp. (English)
35	IUGG/IOC Time Project. Numerical Method of Tsunami Simulation with the Leap-Frog Scheme. 1997. 122 pp. (English)
36	Methodological Guide to Integrated Coastal Zone Management. 1997. 47 pp. (French, English)
37	International Tsunami Survey Team (ITST) Post-Tsunami Survey Field Guide. 2nd Edition. 2014. 120 pp. (English)
38	Guidelines for Vulnerability Mapping of Coastal Zones in the Indian Ocean. 2000. 40 pp. (French, English)
39	Manual on Aquatic Cyanobacteria – A photo guide and a synopsis of their toxicology. 2006. 106 pp. (English)
40	Guidelines for the Study of Shoreline Change in the Western Indian Ocean Region. 2000. 73 pp. (English)

No.	Title
41	Potentially Harmful Marine Microalgae of the Western Indian Ocean Microalgues potentiellement nuisibles de l'océan Indien occidental. 2001. 104 pp. (English/French)
42	Des outils et des hommes pour une gestion intégrée des zones côtières - Guide méthodologique, vol.II/ Steps and Tools Towards Integrated Coastal Area Management – Methodological Guide, Vol. II. 2001. 64 pp. (French, English; Spanish)
43	Black Sea Data Management Guide *(Cancelled)*
44	Submarine Groundwater Discharge in Coastal Areas – Management implications, measurements and effects. 2004. 35 pp. (English)
45	A Reference Guide on the Use of Indicators for Integrated Coastal Management. 2003. 127 pp. (English). *ICAM Dossier No. 1*
46	A Handbook for Measuring the Progress and Outcomes of Integrated Coastal and Ocean Management. 2006. iv + 215 pp. (English). *ICAM Dossier No. 2*
47	TsunamiTeacher – An information and resource toolkit building capacity to respond to tsunamis and mitigate their effects. 2006. DVD (English, Bahasa Indonesia, Bangladesh Bangla, French, Spanish, and Thai)
48	Visions for a Sea Change. Report of the first international workshop on marine spatial planning. 2007. 83 pp. (English). *ICAM Dossier No. 4*
49	Tsunami preparedness. Information guide for disaster planners. 2008. (English, French, Spanish)
50	Hazard Awareness and Risk Mitigation in Integrated Coastal Area Management. 2009. 141 pp. (English). *ICAM Dossier No. 5*
51	IOC Strategic Plan for Oceanographic Data and Information Management (2008–2011). 2008. 46 pp. (English)
52	Tsunami risk assessment and mitigation for the Indian Ocean; knowing your tsunami risk – and what to do about it. 2009. 82 pp. (English)
53	Marine Spatial Planning. A Step-by-step Approach. 2009. 96 pp. (English; Spanish). *ICAM Dossier No. 6*
54	Ocean Data Standards Series:
	Vol. 1: Recommendation to Adopt ISO 3166-1 and 3166-3 Country Codes as the Standard for Identifying Countries in Oceanographic Data Exchange. 2010. 13 pp. (English)
	Vol. 2: Recommendation to adopt ISO 8601:2004 as the standard for the representation of date and time in oceanographic data exchange. 2011. 17 pp. (English)
	Vol.3: Recommendation for a Quality Flag Scheme for the Exchange of Oceanographic and Marine Meteorological Data. 2013. 12 pp. (English)
	Vol. 4: SeaDataNet Controlled Vocabularies for describing Marine and Oceanographic Datasets – A joint Proposal by SeaDataNet and ODIP projects. 2019. 31 pp (English)
55	Microscopic and Molecular Methods for Quantitative Phytoplankton Analysis. 2010. 114 pp. (English)
56	The International Thermodynamic Equation of Seawater—2010: Calculation and Use of Thermodynamic Properties. 2010. 190 pp. (English)
57	Reducing and managing the risk of tsunamis. Guidance for National Civil Protection Agencies and Disaster Management Offices as Part of the Tsunami Early Warning and Mitigation System in the North-eastern Atlantic, the Mediterranean and Connected Seas Region – NEAMTWS. 2011. 74 pp. (English)
58	How to Plan, Conduct, and Evaluate Tsunami Exercises / Directrices para planificar, realizar y evaluar ejercicios sobre tsunamis. 2012. 88 pp. *(*English, Spanish*)*
59	Guide for designing and implementing a plan to monitor toxin-producing microalgae. Second Edition. 2016. 63 pp. (English, Spanish)
60	Global Temperature and Salinity Profile Programme (GTSPP) — Data user's manual, 1st Edition 2012. 2011. 48 pp. (English)

No.	Title
61	Coastal Management Approaches for Sea-level related Hazards: Case-studies and Good Practices. 2012. 45 pp. (English)
62	Guide sur les options d'adaptation en zone côtières à l'attention des décideurs locaux – Aide à la prise de décision pour faire face aux changements côtiers en Afrique de l'Ouest / A Guide on adaptation options for local decision-makers: guidance for decision making to cope with coastal changes in West Africa / Guia de opções de adaptação a atenção dos decisores locais: guia para tomada de decisões de forma a lidar com as mudanças costeiras na Africa Ocidental. 2012. 52 pp. (French, English, Portuguese). *ICAM Dossier No. 7*.
63	The IHO-IOC General Bathymetric Chart of the Oceans (GEBCO) Cook Book. 2012. 221 pp. (English). *Also IHO Publication B-11*
64	Ocean Data Publication Cookbook. 2013. 41 pp. (English)
65	Tsunami Preparedness Civil Protection: Good Practices Guide. 2013. 57 pp. (English)
66	IOC Strategic Plan for Oceanographic data and Information Management (2013-2016). 2013. 54 pp. (English/French/Spanish/Russian)
67	IODE Quality Management Framework for National Oceanographic Data Centres. 2014; revised edition 2019 (English)
68	An Inventory of Toxic and Harmful Microalgae of the World Ocean (in preparation)
69	A Guide to Tsunamis for Hotels: Tsunami Evacuation Procedures (North-eastern Atlantic and the Mediterranean Seas). 2016 (English)
70	A guide to evaluating marine spatial plans. 2014. 96 pp. (English)
71	IOC Communication Strategy for Marine Information Management (2015-2017). 2015
72	How to reduce coastal hazard risk in your community – A step-by-step approach. 2016
73	Guidelines for a Data Management Plan. 2016
74	*Standard Guidelines for the Tsunami Ready Recognition Program*. (in preparation)
75	ICAN (International Coastal Atlas Network) - best practice guide to engage your CWA (Coastal Web Atlas) user community. 2016
76	Plans and Procedures for Tsunami Warning and Emergency Management – Guidance for countries in strengthening tsunami warning and emergency response through the development of Plans and Standard Operating Procedures for their warning and emergency management authorities. 2017
77	IOC Strategic Plan for Data and Information Management (2017-2021). 2017
78	Harmful Algal Blooms (HABs) and Desalination: A Guide to Impacts, Monitoring and Management. 2017
79	IOC Communication and Outreach Strategy for Data and Information Management (2017-2019). 2017
80	Ocean Literacy for All – A toolkit. 2017
81	Procedures for Proposing and Evaluating IODE Projects and Activities. 2018
82	Preparing for community tsunami evacuations: From Inundation to Evacuation Maps, Response Plans, and Exercises (English and Spanish) and Supplement 1 and 2 (English only), 2020.
83	Quality Control of in situ Sea Level Observations: A Review and Progress towards Automated Quality Control, Vol. 1. (English only), 2020.