ESEAS Standards for Sea Level Data Quality Control

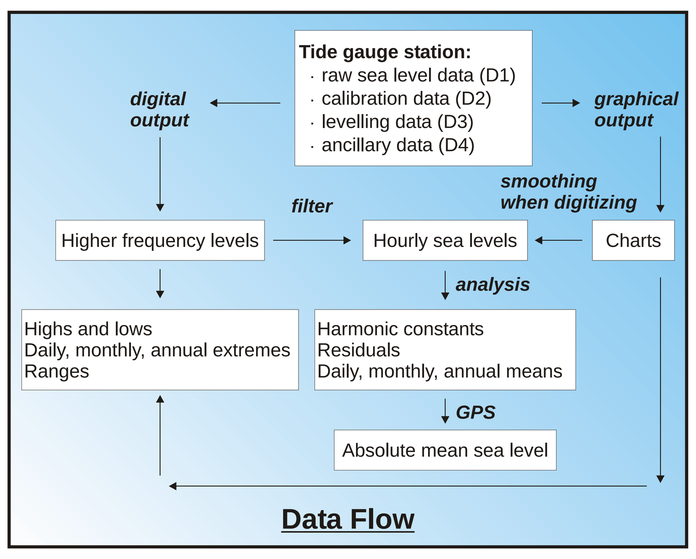
Elizabeth Bradshaw (elizb@bodc.ac.uk)[1], Maria Jesus Garcia (Mjesus.garcia@md.ieo.es)[2], Begona Perez Gomez (bego@puertos.es)[3], Fabio Raicich (fabio.raicich@ts.ismar.cnr.it)[4], Lesley Rickards (ljr@bodc.ac.uk)[1]

[1]British Oceanographic Data Centre, Joseph Proudman Building, 6 Brownlow Street, Liverpool, L3 5DA, UK
[2]Instituto Español de Oceanografía, Corazón de Maria, 8, E-28002 Madrid, Spain
[3]Puerto del Estado, Área de Medio Físico, Avda. Del Partenón, 10, E- 28042 Madrid, Spain
[4]Consiglio Nazionale delle Ricerche, Istituto di Scienze Marine Sezione di Trieste, Viale Romolo Gessi, 2, I-34123 Trieste, Italy

The manual, 'Quality Control of Sea Level Observations', can be downloaded from the ESEAS-RI website,
http://www.eseas.org/eseas-ri/deliverables/d1.2/

1. Introduction

Task 1.2 of ESEAS-RI Work Package 1 was to specify standard routines for the quality control of tide gauge data, with descriptions for tests based on multi-stations.
To fulfil the deliverable, a draft document has been produced and will be outlined here.
One of the objectives of ESEAS and ESEAS-RI is to provide standardised access to a considerable fraction of the European tide gauge data set. It was recognised that a common set of procedures should be adopted for the quality-control of tide gauge data as at the moment most tide gauge authorities tend to use their own methods.
These procedures draw on existing documents e.g. IOC Manuals, the EOSS Position Paper.
Quality control also extends to other factors as well as the data. The documentation and checking of metadata is essential.



2. Metadata (and checks on metadata)

Additional information (metadata) is needed for quality controlling and archiving and is essential when data are exchanged.

- Country (ISO country code)
- ESEAS identifier = PSMSL identifier
- Organization responsible for data collection and processing
- Originator's identifiers for series (e.g. site name and year)
- Geographic location (e.g. latitude, longitude, coordinate system)
- Datum information
- Instrument type
- Description of the installation
- Date and time of the start of the data series (UT)
- Date and time of the end of the data series (UT)
- Time interval
- Quality control status of the series: L1, L2 or not QC
- Data policy
- Available parameters (sea level, atmospheric pressure, wind speed, etc.) and their units

3. Site history and documentation

The documentation of datum information (benchmark relationships), diagrams, maps and other metadata helps to integrate data into a regional or global data set.

Site
- Brief description of location of tide gauge (including maps, photos)
- Description of tide gauge benchmarks (including maps, photos)
- Datum relationships
- Datum history

Data sampling/processing
Brief description of processing procedures used to obtain final data values including:
- Sampling scheme e.g. continuous recording, instantaneous, averaged
- Interval between samples and duration of individual samples (raw data)
- Number of raw data samples
- Nominal interval of processed data
- Gaps in the data record
- Timing and/or datum corrections applied
- De-spiking/smoothing/interpolating methods and editing procedures

Instrument
- Instrument description, manufacturer, model, principle of measurement, method of recording - refer to publication or briefly describe
- Instrument modifications and their effect on the data
- Method and times of calibration, to include calibration factors
- Frequency of cleaning, control of biological fouling
- Operational history
- Pertinent instrument characteristics; for example, for a conventional stilling well, information should include well diameter, orifice depth below mean water level and orifice height above sea bed; for a bubbler gauge - tube length, tube diameter, orifice diameter, density value used to convert to elevation, acceleration due to gravity and the formula used to compensate for tube length.

Report on data quality
- Report any quality control applied by the data supplier to the data set (manual or automatic) or any additional information of use to secondary users which may have affected the data or have a bearing on its subsequent use (e.g. effects of sea state, fouling, etc.).

All quality control procedures applied to a data set should be fully documented. In addition, all quality control applied to a data set should accompany that data set. All problems and resulting resolutions will also be documented. A history record will be produced detailing any data changes (including dates of the changes) made.

Newlyn Tide Gauge

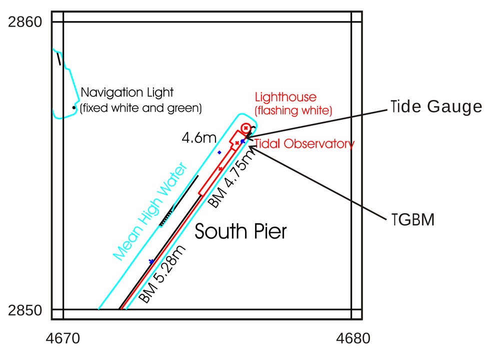
Latitude : 50° 06' 10.8" N   Longitude : 05° 32' 34.2" W

Grid Reference : SW 4676 2856

Instrument type : Data acquisition system with a full tide, a mid-tide bubbler gauge and a potentiometer attached to a Munro float gauge installed.

Site of Gauge:

The Tidal Observatory is located at the end of South Pier, Newlyn, next to the lighthouse, and the measuring points are located on the seaward side of the pier behind the lighthouse.



Benchmarks and Benchmark relationships:

Benchmark	Grid Reference	Description
TGBM	SW 4677 2856	Brass bolt in the floor of the recorder hut.
Aux1	SW 4673 2851	Flush Bracket 1565 on wall S pier NW face 17.8m SW.
Aux2	SW 4659 2841	F Bracket 1520 wall SE side of S Pier Rd NW face

TGZ = Admiralty Chart Datum (ACD)
TGZ = 3.05m below Ordnance Datum Newlyn (ODN)
TGZ = 7.801m below TGBM

Ordnance Datum Newlyn (ODN) is based on mean sea level at Newlyn between 1915 and 1921 (inclusive).

Datum information : All data are to Admiralty Chart Datum (ACD).

Levelling information : The site was last levelled on 08/07/1997.

T.G.I. visits to site : Day 113 Float gauge recalibrated.
Day 317 Mid-tide channel connected.

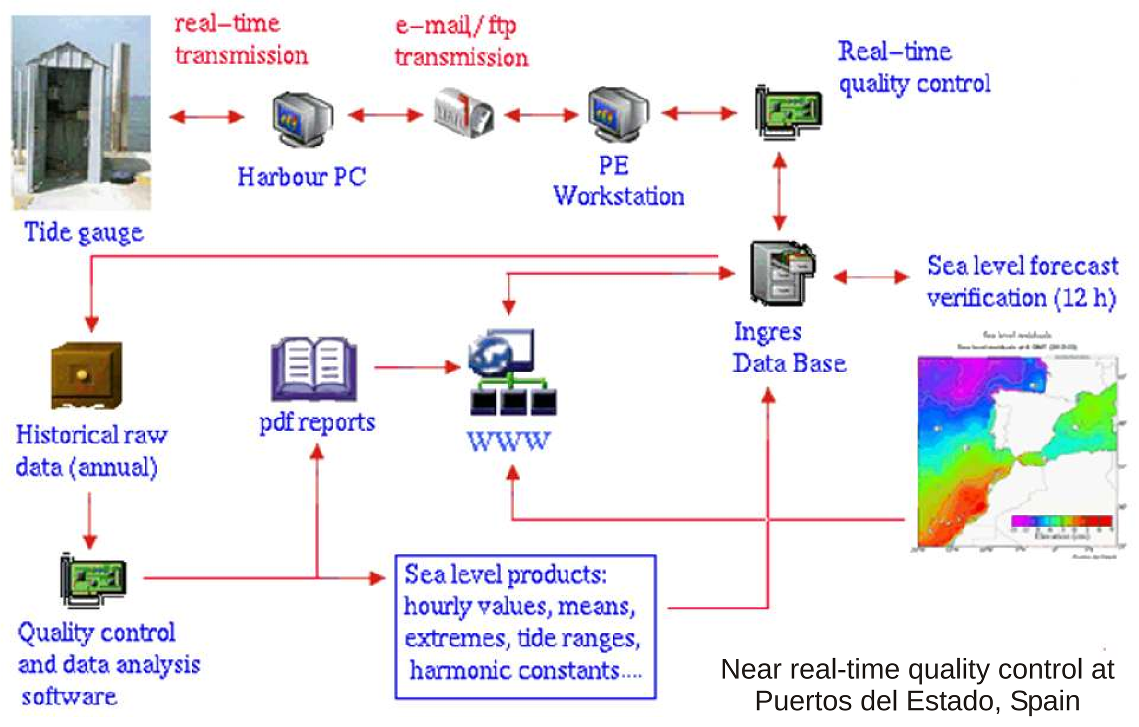
Data quality

Only the primary channel (parameter name ASLVBG02) has been screened and quality controlled, so the results presented here are from that channel only.

CI (%)	Sample interval	Missing data	Suspect data
90	15 minutes	066-100, 114-115, 317	None

Example of document to accompany tide gauge data

4. Quality control and quality assurance



Near real-time quality control at Puertos del Estado, Spain

4.1. Quality flags

The process of quality control consists of performing various checks and depending on the results may change the value or leave the data unchanged. In both cases the data should be flagged according to pre-defined quality control codes. Within ESEAS a standard set of quality flags should be agreed.

If data values are altered as a result of quality control practices, it is strongly advised that the original value be preserved with the data.

ESEAS FLAGS	Mandatory
0 - no quality control	No
1 - correct value	Yes
2 - interpolated value	Yes
3- doubtful value	No
4- isolated spike or wrong value	Yes
5 - correct but extreme value	No
6 - reference change detected	No
7- constant values for more than a defined time interval	No
8 - out of range	No
9- missing value	Yes

4.2.Automatic quality control procedures

If data are inspected daily, either manually or automatically, then faults can be rapidly attended to and fixed.
Raw data quality control (hourly or more frequent i.e. 5, 6, 10, 15 minutes):

- checks of strange characters
- checks of date and time
- out of range and spikes detection
- constant values detection (stability test)
- doubtful values (ex: target 00 for acoustic sensors)

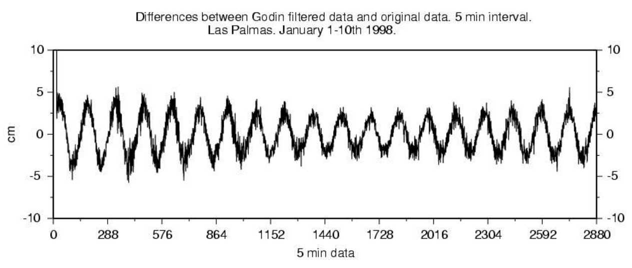
The result is the original raw data flagged as explained above, and a new corrected time series with flags, regular sampling and gaps lower than a certain interval interpolated.

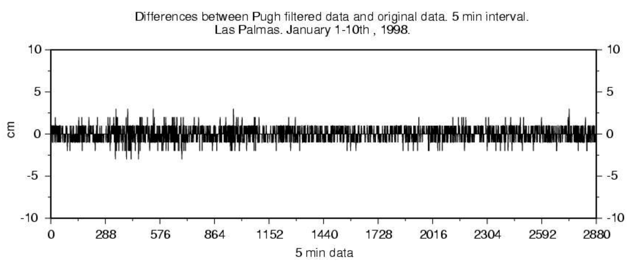
4.3. Manual or 'scientific' quality control

The final goal of quality control is to detect and, if possible, correct errors, in order to minimise the related loss of information.

- For analogue records, the chart conditions and mounting must be checked, as well as the continuity between subsequent charts.
- Produce a tidal analysis (compute 'fresh' tidal constants, don't rely on historical values). This can be performed by means of software like those developed by the University of Hawaii, the PSMSL/Proudman Oceanographic Laboratory and the Australian National Tidal Facility. Tidal constants used in tide predictions should never be mixed between different packages.
- Inspect the recorded data and residuals for unphysical values and instrumental faults such as timing errors, datum shifts, spikes/outliers, gaps, etc.
- Check against records of a nearby site, to see if the suspect values are due to a tide gauge fault or to meteorological conditions.
- In case of a fault data should be corrected or interpolated, otherwise must be kept as they are, taking note of the event.
- If possible, more than one instrument should be operated at the same site, in order to fill gaps by direct comparison. A mechanical device with analogue record still represents an invaluable support system.

Filtering
Standardised filtering techniques will ensure that data supplied to the global sea level databanks (e.g. PSMSL and IHO) are consistent
PE and IEO performed tests comparing a filter described by Pugh (1987) for 5, 10 or 15 minutes interval data, and a filter discussed by Godin (1972) that can be applied to any time series at any interval. The data resulting from applying the second procedure gave a smoothed tidal amplitude. In fact, FOREMAN in its tidal analysis software has a procedure to recover the amplitude.

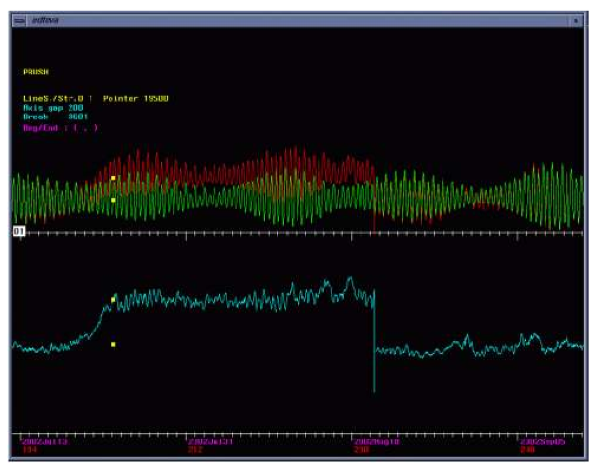




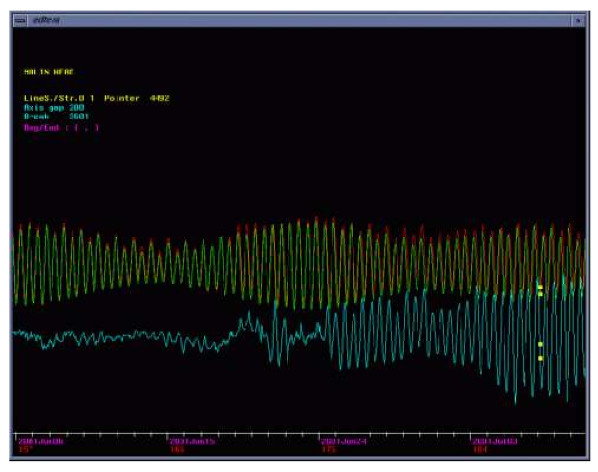
Stability: Reference changes
Most shifts are readily identifiable in hourly residual plots, plots of the daily and monthly data and in plots of differences of daily or monthly values with nearby tide stations or with redundant sensors at the site in question. If a shift is identified, the proper means of correction is through the analysis of the tide staff readings and the corresponding tide gauge values. The reference change value should then be added over the corresponding period.

Shift and drift time
A shift or drift in the time (caused either by operator error or clock malfunction) can be detected by visualization of the residuals or correlation between observed data and predictive or neighbourhood station data. The data will then need to be moved according to the lag time (if constant) or interpolated.

Stability: Shift and drift data
Changes in the instrumentation and in the environment surrounding the station can result in sharp discontinuity (shift datum) or can lead to trend (drift datum). Checks should be performed to obtain a unique reference. The shift and drift data error can be detected by visualization of the mean sea level of one station and better as differences of two stations.
To correct this, add the shift value or de-trend the drift value and document the action.
The vectorial mean harmonic constants can be calculated, and using also existing annual ones, the variability of each constant from year to year can be computed and plotted. For some applications those with too much variability are just discarded for prediction.

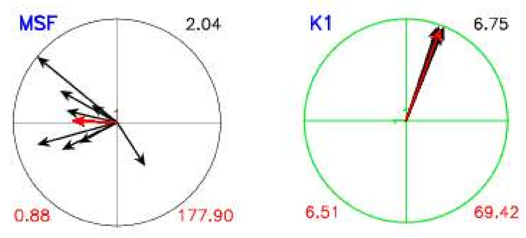


Screening software showing datum shift



Screening software showing timing error

Interpolations
During quality control, the technique to replace the spikes for an appropriate value is linear interpolation.



Plots of vectorial mean harmonic constants

Correlation
Correlations 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 this is a valuable tool for detecting problems. The correlation analysis is also useful for filling gaps.

4.4. Final quality control and products

Annual review, calculation of statistics

In order to validate the data, information on general behaviour of the station or grouped into oceanographic and/or climatic region needs to be calculate from historical data. In particular some statistics should be performed to obtain broad ranges means and standard deviation to perform some quality checks.

- upper and lower limits or historical extremes (for range check).
- tidal and observed sea level ranges
- extremes, mean and standard deviation of hourly values, meteorological residuals, ranges or mean sea levels
- tables of monthly and annual extremes
- density function for hourly values, tide predictions and residuals

(*)Broad ranges and standard deviation tables for the sea level and related parameters have to be prepared.

4.5. Anomalies and trends

Test for non-homogeneity

Standard Normal Homogeneity Test
The SNHT gives the points where an inhomogeneity exists and provides information about the probable break magnitude.

Test for trends

EOF Analysis
An EOF analysis over a set of tide gauge records (2 or more) is an available method to estimate the trends and to detect possible errors. The stations must be on the same coastal area because in this case the sea response to the forces could be the same along the coast.

Mann-Kendall test for trends
This non-parametric test can be used to study the increases and decreases and determine temporal trends in climatic data sets.

5. ESEAS Data Exchange format

A standard ESEAS format has been agreed. A format description can be found within the 'Quality Control of Sea Level Observations' manual.

In task 1.3 of ESEAS-RI Work Package 1 (Web interface to data centres) the data format and comprehensive information required to accompany data has been developed.

6. ESEAS Software Package for processing and quality control

ESEAS-RI may, if necessary, implement new software. Currently available software includes:

- The University of Hawaii has produced a package for sea level data quality-control purposes.
- The PSMSL/Proudman Oceanographic Laboratory, has produced a package called TASK-2000 (Tidal Analysis Software Kit) which is based in the TIRA etc. programmes used at POL for many years.
- The Australian National Tidal Facility, which has own package derived originally from POL software.

Within the ESEAS-RI project, two additional software packages are available for the sea level community:

- Puertos del Estado sea level software, which has been developed in PE for Unix systems, mainly automatic, and based on the original UHSLC software (same formats).
- IOLR (Israel Oceanographic and Limnological Research) sea level software for Windows.