Chapter 11

European Sea Level Monitoring: Implementation of ESEAS Quality Control

María Jesús García,Instituto Español de Oceanografía, Corazón de María, 8, E-28002 Madrid, Spain
Begoña Pérez Gómez,Puertos del Estado, Área de Medio Físico y Tecnología de las Infraestructuras, Avda Partenón, 10, E-28042 Madrid, Spain
Fabio Raicich,Consiglio Nazionale delle Ricerche, Istituto di Scienze Marine, Viale Romolo Gessi, 2, I-34123, Trieste, Italy
Lesley Rickards, Elizabeth Bradshaw, British Oceanographic Data Centre, Joseph Proudman Building, 6 Brownlow St., Liverpool, L3 5DA, UK
Hans-Peter Plag, Nevada Bureau of Mines and Geology and Seismological Laboratory, University of Nevada, Reno, Mailstop 178, NV89557, USA
Xiuhua Zhang, Bente Lilja Bye, Espen Isaksen,Geodesi, Norwegian Mapping Authority, 3507 Hønefoss, Norway

Abstract. One of the objectives of the European Sea-Level Service (ESEAS) is to provide access to quality controlled European tide gauge data for research and other uses. The ESEAS quality control is based on a common set of procedures for the quality control of observed sea level data, which to a great extent can be referred to those specified by the Intergovernmental Oceanographic Commission for the Global Sea Level Observing System. Quality control also extends to other information such as documentation of datum information, metadata, exchange format, application classification, and levels of quality control. In relation to tsunami-enabled tide gauges, quality control is related to the availability of data in real-time and the use of automatic control and analysis procedures.

Keywords. Sea level monitoring, tide gauge, distribution of data, standardization, quality control

1 Introduction

The European Sea-Level Service (ESEAS) is an international collaboration of organizations in 23 countries. It has brought together the formerly scattered sea level observation and research community in Europe as a major research infrastructure for all aspects related to sea-level, in the fields of climate change research, natural hazards or marine research. One of the main objectives of the ESEAS is to provide access to quality controlled European tide gauge data. Most tide gauge authorities have developed their own methods of quality control and for the ESEAS, a common standardized set of procedures has to be adopted. Based on existing documents such as the IOC manuals (UNESCO, 2002, and the references therein) and ESEAS documents (Plag et al., 2000), these standards were defined within the European Sea Level Service-Research Infrastructure (ESEAS-RI) project, funded by the European Commission from November 2002 to October 2005.

The ESEAS Observing Sites (EOS) are primarily tide gauges providing continuous measurements of the sea level relative to a well monitored benchmark on land (ESEAS Governing Board, 2001). Some of these tide gauges are co-located with geodetic techniques such as continuous Global Positioning System (CGPS), and in some cases episodic absolute gravity measurements are performed. For an EOS committed to the ESEAS, operation has to comply with the accepted standards for each of the applications defined and classified by ESEAS (2004a). At present, there are more than 200 tide gauges in the ESEAS observing network distributed along thousands of kilometres of the European coast.

Two levels of quality control depending on the delivery time line are defined, namely L1: quality control for low-latency products, and L2: full quality control and analysis for delayed mode products. L1 is less detailed than the one applied to delayed mode or historical time series, and is tailored for the operational use of sea level data (e.g. storm surges, tsunami warning), where automatic procedures, short sampling intervals and rapid transmission are crucial. For the L2 or delayed mode quality control, the use of additional information is mandatory: documentation of datum information, diagrams, maps and other meta-data have to be provided. However, there has been little standardization of the type of information. The ESEAS has agreed on common procedures across the participating organizations.

The design of a unique ESEAS web interface is in progress (www.eseas.org). This allows easy access to both the near-real time and historical quality controlled data. The interface provides a single point of access for users, who previously needed to contact individual institutions of different countries in order to get information and data, if available, about sea level.

2 Data Requirements

From each tide gauge station three basic data sets should ideally be available, namely (1) raw sea level data in digital or graphical form, (2) the calibration data (normally manual data obtained at the site during each calibration), and (3) levelling information. Sometimes ancillary data are also present. Depending on the acquisition system, the actual sea level data are obtained either by digitizing graphical records of analog systems, or from a sensors providing digital output. In the latter case raw sea level data are obtained at a particular sampling interval (at present between 5 and 10 minutes, normally not lower). These raw higher frequency data are processed to L1 quality control, in order to be available for operational purposes. Hourly values are obtained by means of an adequate filter.

Additional information (metadata) is needed not only for quality control and archiving, but also for exchanging data or integration of them into a regional or global data set. A wide variety of metadata 'standards' (e.g. FGDC, Dublin Core, ISO19115) are in regular use and metadata recommendations are also available from IOC Manuals and Guides, GE-TADE (IOC/IODE), and various other projects. The ISO standard ISO19115 for geo-spatially referenced data has recently been ratified, thus the metadata for ESEAS will be compliant with this standard.

The ESEAS Quality Control document (ESEAS, 2004b) includes a list of the information that should be stored with the data. The metadata accompanying each data series must be sufficient to ensure that the data are adequately qualified and may therefore be used with confidence. Such documentation should be stored alongside the data, and where applicable, should describe the site, the data sampling and processing and the instruments used.

3 Quality Control and Quality Assurance

As for any measured physical value, the errors that can arise in the sea level data and related parameters can be random or systematic. Examples of the first type are electronic noise, problems in the communications, sensor calibrations, etc. Systematic errors would be, for example, inhomogeneities due to changes in observational practices, or a change of instrumentation. Changes of instrument location can result in sharp discontinuities in the sea level data. Changes in the environment of the station, such as harbour constructions, land movements, etc., can produce trends in the data or changes in the tide parameters.

For the L1 level (low latency), the initial process of quality control consists of performing various checks of the original series and then flagging of suspicious values according to ESEAS quality control codes. Near-real time data applications require the development of automatic procedures that perform the basic quality control also in near-real time. This automatic procedure consists basically of the following steps applied to the original high frequency raw data:

- checks for strange characters
- checks of date and time
- gap detection
- out of range and spikes detection
  - constant values detection (stability test)
  - doubtful values

An adequate algorithm for detection of spikes has to be used, which leaves real phenomena, like tsunamis and seiches, unflagged. At this stage normally ancillary data or data from nearby stations are not used as the time of processing would become too long and ineffective. The result of the L1 quality control is the original raw data flagged, and a new corrected time series with flags, regular sampling and gaps lower than a certain interval interpolated.

As part of the scientific or delayed mode quality control (L2), a more detailed processing of sea level data is performed, applied to longer time series (typically one year) that include not only the steps described for L1, but also the filtering to hourly values, computation of annual harmonic constants, residuals, extremes and means. The examination of the quality of these products is crucial for the detection of problems and malfunction in the tide gauge. The primary quality control of sea level is based on the inspection of both recorded data and meteorological residuals.

Raw data are normally registered at time intervals between 1 minute and 1 hour, the most common being 5, 6 and 10 minutes. In regions where seiches occur frequently, or where phenomena such as tsunamis should be detected, sea level data are registered at 1minute interval or even less. Apart from keeping higher frequency signals for other purposes, it is always necessary to obtain filtered hourly values for further sea level data processing. The filtering process will eliminate higher frequencies dependent 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. Within the ESEAS the recommendation is to use filters with amplitude response similar to the one of Pugh or Butterworth filters, in order to avoid reduction in tidal ranges and extremes just due to filtering. A study of the importance of higher order harmonics for a particular station may indicate the need for designing a filter, which does not attenuate other high frequencies.

A common procedure is to compute the harmonic constants for each year of observed data. Some problems in the data series, such as clock errors, will introduce temporal variation in the normally stable harmonics. An inspection of the variation through the years of the harmonic constants is interesting both for detecting problems and also for having information about changes at the station. For example, changes in the configuration of a harbour can affect the tide parameters.

The inspection of meteorological residuals is a very useful tool for the quality control process. All fundamental types of errors that a sea level series can present (e.g. clock malfunction, reference changes, spikes) are easily detected in the residual plot. 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.

Correlations can be computed both between data from different stations or sensors and between different parameters at the same station (wind, atmospheric pressure, etc), and this is a valuable tool for detecting problems. Study of the lag correlation between observed and predicted data is used to detect clock malfunctions.

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

- 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,

For historical data, the lack of sufficient metadata often complicates the quality control. This may particularly affect the control of datum stability. In this case, some additional checks should be performed to obtain a unique reference. A standard procedure is to work with several sea level series from nearby stations.

Different algorithms described in literature can help to detect discontinuities or reference problems in historical data or hourly time series. The Standard Normal Homogeneity Test (SNHT, Alexanderson, 1986) gives the points where an inhomogeneity exists and provides information about the probable break magnitude. An EOF (Empirical Orthogonal Function) analysis over a set of tide gauge records (2 or more) is an available method to estimate the trends and to detect possible errors, apart from other applications. According to Sneyers (1975; 1992), the non-parametric Mann-Kendall test for trends can be used to study the increases and decreases in climatic time series.

4 Software Packages for Sea Level Quality Control and Processing

Within the ESEAS, the development of a standardized ESEAS sea level data quality control and management software has been proposed. Most sea level data collecting organizations within ESEAS have developed their own software to validate incoming data in varied formats and media that are specific to their requirements. However, it would benefit the community if there would be a platform independent package available. An on-line tidal analysis facility has been made available on the UK National Tidal and Sea Level Facility website (www.pol.ac.uk/ntslf), and the Israel Oceanographic and Limnological Research (IOLR) Institute has produced a Windows based sea level package. These, together with other software packages currently in use provide the basis of an ESEAS package.

5 ESEAS Data Exchange

The ESEAS has agreed on an exchange format, including a set of quality flags. The format is a simple 'spreadsheet' style similar to that in use in many participating organisations and has been defined taking into account international guidelines developed by working groups within the IOC's International Data and Information Exchange Committee. The format includes appropriate ancillary information. Via the website, users can search for data by date, region, country and site. The retrieved data contain information about the quality control of the data, thus allowing each user to decide for what applications the data are suitable for. The basic principles of the web portal are described in Plag (2004). The web portal is designed as a distributed system with the tide gauge data being stored at each ESEAS Operational Centre (EOC). A database is used to store metadata for the tide gauge data as well as details of the EOCs and the EOSs according to a predefined structure. An underlying principle in developing the web portal is that the work required by each EOC should be kept at a minimal level. When a user submits a request for data, the web portal will automatically retrieve the data from the respective EOC, thus ensuring that the data are up-to-date. The standardized quality control described above ensures that the quality of the products from the web portal is as far as possible independent of the EOC delivering the data.

Acknowledgments

The authors would like to thank the many colleagues in the ESEAS and the ESEAS-RI project for the discussions of the ideas reported here. The ESEAS-RI project was supported by the European Commission under the contract number EVR1-CT-2002-40025.

References

Alexanderson, H. (1986). A Homogeneity Test Applied to Precipitation Data, *Journal of Climatology*, 6, pp661-675.

ESEAS (2004a). Classification of ESEAS Operational Sites for Applications A- D, http://www.eseas.org/eseas-ri/deliverables/d1.3

ESEAS (2004b). Standards for Quality Control of Tide Gauge Observations, Deliverable D1.2. http://www.eseas.org/eseas-ri/deliverables/d1.2

ESEAS Governing Board (2001). European Sea Level Service, Terms of Reference, http://www.eseas.org/websitemap.php

Foreman, M.G.G. (1977). Manual for Tidal Heights Analysis and Prediction, Pacific Marine Science Report 77-10, p100.

Godin, G. (1972). The Analysis of Tides, Liverpool University Press, Liverpool, p264.

Plag, H.-P., Axe, P., Knudsen, P. Richter, B. and Verstraeten, J.,(2000). European Sea Level Observing Systems (EOSS) Status and Future Developments, European Commission, Cost Action 40, EUR 19682.

Plag, H.-P. (2004). The ESEAS Data Portal: Principal Considerations, In Holgate, S. and Aarup, T. (ed.): *Workshop on New Technical Developments in Sea and Land Level Observing Systems*, IOC Workshop Report No.193, pp108-113.

Pugh, D.T. (1987). Tides, Surges and Mean Sea-Level, Chichester, John Wiley and Sons, p472.

Sneyers, R. (1975). Sur l'annalyse statistique des séries d'observations, Nº 415 WMO, Ginebra.

Sneyers, R. (1992). On the Use of Statistical Analysis for Objective Determination of Climate Change. *Meteorol. Zeitschrift*, No. 1, pp247-256.

UNESCO (2002). Manual on Sea Level Measurement and Interpretation, Volume III: Reappraisals and Recommendations as of the Year 2000, IOC Manuals and Guides, No. 14, p55.