This paper was presented at the Ocean Data Symposium, Dublin, Ireland (October 1997)

Developments in sea level data management and exchange

Lesley Rickards
British Oceanographic Data Centre
Proudman Oceanographic Laboratory
Bidston Observatory
Birkenhead, Merseyside, L43 7RA, UK

Bernard Kilonsky
University of Hawaii Sea Level Centre
Department of Oceanography
University of Hawaii
Honolulu, HI 96822, USA

Abstract

Responsibility for data management of the *in-situ* sea level data set for the World Ocean Circulation Experiment (WOCE) is jointly vested in the University of Hawaii Sea Level Centre (UHSLC) and the British Oceanographic Data Centre (BODC). The two centres perform complementary activities, with UHSLC acting as the 'fast delivery' centre, providing rapid access to data for those involved in satellite altimetry, and BODC as the 'delayed mode' centre, providing the fully quality controlled comprehensive data set at the end of the project.

The activities of the two centres will be described, including the quality assurance procedures currently in place and those under development, formats for data storage and exchange, and the metadata to be stored alongside the data. In line with the plans to set up a distributed sea level information network, as described in the revised IOC Global Sea Level Observing System (GLOSS) Implementation Plan, we will review progress and describe the developments underway to make sea level data more easily accessible to the scientific community.

Introduction

The last several years have seen the continued development of *in-situ* sea level measurements as an important oceanographic data set. Much of this progress came about under the World Ocean Circulation Experiment (WOCE) programme through the efforts of the British Oceanographic Data Centre (BODC) and the University of Hawaii Sea level Centre (UHSLC).

WOCE, as a component of the World Climate Research Programme (WCRP), is a cooperative scientific effort by more than 30 nations studying the role of the ocean circulation in the earth's climate system. WOCE has evolved into two overlapping phases. The project was planned to begin in 1990 with observations planned through

1994 and a further five years of analysis to the year 2000. However, delays in launching several satellites and cruises led to an extension of the field phase to 1997. The WOCE field programme has focused on measurements of the deep oceans and the development of new technologies for ocean exploration.

The WOCE analysis and synthesis efforts now scheduled to continue at least through 2002 will focus on developing new methods for handling and analyzing data and designing improved models for use in climate prediction. As knowledge of the ocean circulation is necessary for predicting climate variability, the WOCE implementation plan devised a decentralized data management system to assure that important oceanographic data were assembled, quality controlled and archived. The system utilizes the expertise of scientists familiar with specific oceanographic data types at research institutes and national data centres to produce the highest data quality and documentation.

Under the WOCE data management system data flows from the data collectors, usually principle investigators, to Data Assembly Centres (DAC) and Special Analysis Centres (SAC). The data are quality assured and then distributed to end users and archived. The DACs and SACs ensure the direct involvement of research groups in the management of WOCE data set. They are directed by scientists and accomplish the basic assembly and quality control of the WOCE data, generate data products, and perform data analysis functions including the production and distribution of derived data sets. Because of the distributed structure of the data management, a WOCE Data Information Unit was developed to serve as a source of information on the WOCE field programme by tracking all data collection, processing and archiving activities, and arrangements were made with the World Data Center-A for Oceanography (WDCA) to serve as the permanent WOCE data archive.

While the WOCE field programme has assembled a multitude of surface, subsurface, and satellite based data sets, the rest of this paper will focus on the WOCE *in-situ* sea level DACs and their data sets. We will first provide some background on the two centres, the BODC and UHSLC, then review their efforts collecting, processing and distributing the WOCE data, and finally we will examine the current state of the data base and discuss what the future might hold for *in-situ* sea level data management and exchange.

Historical perspective

The establishment of WOCE DACs at the BODC and the UHSLC was not the first experience these two centres have had with collecting, processing, and distributing *in-situ* sea level data. The Proudman Oceanographic Laboratory (POL), where BODC is located, has been involved with sea level data for many years. The UK permanent tide gauge network was set up as a result of severe flooding along the east coast of England in 1953. The network now comprises approximately 45 gauges related through the national levelling network to Ordnance Datum Newlyn [1]. Data are collected, processed and banked centrally at BODC to provide long time series of reliable and accurate sea levels. The data are used for tidal analysis and prediction, oceanographic research, coastal defence and storm surge warning systems, this latter operating in near-real time. The Tide

Gauge Inspectorate based at POL is responsible for the operation, maintenance and development of the tide gauge network under the guidance of the UK Committee on Tide Gauges. Also co-located at POL is the Permanent Service for Mean Sea Level (PSMSL), the international repository for mean sea level data and information which operates under the auspices of the International Council of Scientific Unions (ICSU).

The BODC itself, has in its remit, the maintenance and development of the UK national oceanographic data bank, which includes sea level data amongst other data types. The BODC has over 20 years of experience in the field of data management for a variety of oceanographic data types.

The BODC was also given responsibility for the sea level data set collected during the Mediterranean Alpine Experiment (MEDALPEX), which took place in 1981-82. Sea level data were requested from seven Mediterranean countries and passed through the usual BODC data banking procedures. In addition to this the data were tidally analyzed and low pass filtered. Obviously to quality control the data this thoroughly was highly desirable, and allowed problems to be solved which might otherwise have gone unnoticed. Once the data set had been compiled and quality controlled a report and magnetic tape containing the data set were produced [2] and lodged with the WDA system.

The UHSLC, under the direction of Klaus Wyrtki, began installing tide gauges in the tropical Pacific Ocean in the early 1970s [3]. The maturation of this activity in the intervening years has resulted in a mostly island based network of gauges in the Pacific and Indian oceans that is presently the largest open ocean sea level network in the world operated by a single group. In 1985 the data collection activity represented by the UHSLC field programme was complemented with the co-location at the UHSLC of the Tropical Ocean-Global Atmosphere (TOGA) Sea Level Centre [4]. This centre collected hourly sea level data from many data originators, of which the UHSLC field programme was a primary one. The TOGA data set was distributed directly from the UHSLC and was also archived in the WDC system and the U.S. National Oceanographic Data Centre (NODC) for additional distribution. The UHSLC also currently operates a Specialized Oceanographic Centre (SOC) for the Integrated Global Ocean Services System (IGOSS) [5]. The IGOSS Sea Level Project in the Pacific is an early, and very successful, example of operational oceanography. Since June 1984 monthly maps of the Pacific sea level topography have been produced without fail. The analysis includes monthly mean sea level data from 93 participating stations in 33 countries throughout the Pacific basin.

When WOCE recognized the need for *in-situ* sea level data, it was only natural to take advantage of the experience that already existed at both BODC and the UHSLC and request that they be established as DACs.

The WOCE *in-situ* sea level centres

In-situ sea level measurements provide time series at fixed locations over periods of many decades. These observations have many applications including the analysis and interpretation of the ocean's dynamics and the monitoring of short-term climatic change. The WOCE implementation plan [6a and b], which placed a heavy reliance on sea surface height measurements derived from satellite-borne altimeters, identified the requirement for in-situ sea level measurements to:

(1) support altimetry using a geographically well-distributed global network of (mostly) island gauges,

(2) compare the observations made during the WOCE period to longer time series in order to evaluate the representativeness of the WOCE time frame, and

(3) estimate geostrophic currents using sets of gauge pairs across straits.

Additionally the WOCE Data Handbook [7] stated that hourly, or more frequent observations are required and, for locations at mid or high latitudes, the tide gauge measurements must be supplemented by sea level atmospheric pressure data. The strategy for tide gauges in WOCE was to take advantage of the existing extensive regional networks such as the TOGA programme and extend them in accordance with the following needs:

(1) to complement altimetric measurements in oceans with sparse island distributions;

(2) to instrument the high latitude Southern Ocean, both to complement altimetric measurements and as an independent measure of variability in a poorly observed region;

(3) to instrument straits and channels (e.g. Drake Passage), through which there is considerable transport, which can be monitored by surface elevation measurements.

Responsibility for the collection and management of the *in-situ* sea level data set for WOCE was jointly vested in the two centres. The UHSLC was established as the 'fast delivery' DAC and tasked with the assembly, quality control and distribution of all sea level data from WOCE gauges delivered by satellite or other near real-time systems. The data were to be made available to investigators in a time frame of 1 - 3 months after data collection. The BODC, as the 'delayed mode' DAC, was to assemble and supply sea level data from the WOCE network to the full extent of quality control. Distribution was to be within 18 - 24 months after data collection. BODC was also tasked to ensure archival of the sea level data as a WOCE data set in the WDC system by the end of the WOCE experiment.

The 'fast delivery' sea level DAC

The creation of the WOCE DAC necessitated a major new initiative for the UHSLC. Before it was established, sea level data were collected, processed, and distributed within 1 - 2 years after the calendar year of the data collection. The WOCE 'fast delivery' DAC, on the other hand, was to provide information needed to check the altimeter data against the more traditional and well-understood sea level data from the tide gauges. The altimetry data were available within a month or so of collection. Thus, the UHSLC had to process data from a globally-distributed set of stations and make the *in-situ* sea level data available to users on a comparable time scale. The turnaround time for this data set is much faster than for the TOGA data set, and the geographical extent of the data set exceeds that of the TOGA and IGOSS data set. Fortunately, the UHSLC had access to a large fraction of the open ocean gauges that were capable of reporting sea level data quickly enough to meet the near real-time requirement. The UHSLC also had experience with distributing data and data products on the required time scale via the IGOSS sea level project.

The UHSLC *in-situ* sea level stations employ a robust design that emphasizes redundancy of measurements including an automated switch that produces reference level information [8]. These stations provide long-term sea level monitoring accurately related to a datum at the millimetre level using inexpensive float-operated gauges. They upload data to the UHSLC via the NOAA's Geostationary Operational Environmental Satellite (GOES) Data Collection System (DCS), Japan's Geostationary Meteorological Satellite (GMS) DCS, and the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) Meteorological Satellite (METEOSAT) DCS [9], and provide the nucleus of the WOCE 'fast delivery' data set. The UHSLC has also expanded upon the contacts made with other national agencies contributing sea level data to the TOGA and IGOSS projects. Through these activities, we have significantly augmented our collection of near-real time sea level data.

The 'delayed-mode' sea level DAC

BODC's role as the 'delayed-mode' DAC is to assemble, distribute and supply sea level data to the full extent of quality control possible covering all of the gauges in the network. WOCE requirements are that the elevations should be accurate to 1 cm, the timing to 2 minutes and the atmospheric pressure measurements to 1 mbar. Operation of the 'delayed mode' sea level DAC meant an expansion of BODC's previous sea level activities to complement the 'fast delivery' centre activities described above. Data collected by the UHSLC as part of the TOGA data set provided a significant element of the initial data set, but contacts have also been initiated with some 25 organizations around the globe requesting their sea level data for WOCE. Data are requested annually, with atmospheric pressure data requested in addition to sea level, although these have not proved easy to obtain. Special effort has gone into obtaining data from gauges which not readily accessible in near-real-time.

In addition to the sea level data collected from standard tide gauges, BODC also has the responsibility for data collected by bottom pressure recorders and inverted echo sounders. Much of the contribution for this part of the data set comes from the POL ACCLAIM (Antarctic Circumpolar Current Levels by Altimetry & Island Measurements) network in the South Atlantic. The BODC database is flexible so that data with more than one parameter, or indeed, with different parameters can easily be incorporated.

Data Collection and quality assurance

When data arrive at the 'delayed-mode' DAC on magnetic media they are checked for readability and the data supplier informed of any problems; if necessary the data can be resubmitted, although this is a rare occurrence. Backup copies are made of the incoming data and they are transferred to BODC's internal (binary) format to allow checking and quality control of the data. Some software is written for each new format received, to interface with the main checking software. The data undergoes many checks at this stage to ensure that the format is consistent throughout and that the parameters and their units are properly defined. Most data are received in units of metres/centimetres/millimetres but some data are received in millibars where, for example, bottom pressure recorders have been used.

The sea level and accompanying data are usually supplied annually and, more importantly, each submission usually comprises at least 1 year of data from each site. If data of shorter duration are received they are quality controlled when a year of data is available unless, of course, the tide gauge is expected to be non-operational for a long period. There are several reasons for this approach, the most important being that it is easier to detect and correct datum shifts using a longer time series and the effort required to check 1 year of data is similar to that required for 1 month of data. Most often the data are supplied as hourly values, but there is no restriction on this. Where possible gaps in the data will have been filled from redundant sensors. However this is not always possible and gaps in the data are flagged as null data and documented.

There are three main components to the sea level data quality assurance: flagging obvious wrong values, correcting timing drifts and maintaining reference level stability. Quality control is carried out with the aid of high speed graphics workstations, utilizing the in-house EDTEVA software package, which has been written specifically for sea level data processing and quality control.

Sea level data values are checked for spikes and physically unreasonable values. Tidal predictions can also be displayed. These may be generated from the POL tidal constituents data bank or by tidally analyzing the data to produce constituents and thence the predictions. The former method is used for the UK network and the latter for the WOCE data. The residuals produced from tidal analysis are screened to check for datum shifts, timing problems and other errors.

For the data received for WOCE, tidal analysis is performed on each year of data. The Doodson type of harmonic analysis, which has been in use for many years at POL, is used. Constituents can then be compared with those obtained from previous years of data and also with values in the International Hydrographic Bureau (IHB) or POL constituents data bank. The harmonic constituents produced by the tidal analysis of one year's data can be used to predict the tides for other years, if necessary.

The screening software allows the sea level data values to be displayed as a stacked monthly plot or as a simple time series (i.e. parameters plotted against a single time axis).

In this latter presentation all the parameters measured can be displayed, or a subset can be chosen. Observations and predictions can be plotted on the same horizontal axis if required. Adjacent stations are compared to check out unusual signals. This is done by grouping the series together prior to using the package. The data series can then be viewed together, one plotted over the other as required. In addition, any other parameters, for example, atmospheric pressure or sea surface temperature, are also displayed and examined. Daily and monthly means can also be calculated.

Problems with the time channel will show up in the residuals obtained through the tidal analysis. Where shifts in the datum level are suspected, these are referred back to the data supplier and necessary adjustments made to the data. If shifts in the datum level cannot be resolved, they will be documented and this information stored with the data. Tide gauge records are split into separate files if the reference levels on either side of a shift are not linked by levelling to the same benchmarks.

A quality assessment is carried out on the data. This will indicate the proportion of available data that do not contain questionable fluctuations in the residuals. Each case must be carefully checked to determine whether the fluctuation is a natural event, an indication of mechanical problems with the tide gauge, or the result of unreliable predicted tides. All such features will be noted in the documentation stored with the data.

The qualifying information accompanying the data is also checked (Table 1). This includes assembling datum information; the local datum should be recoverable through a set of permanent and auxiliary benchmarks. Details of the datum and its method of determination is also required as are benchmark and levelling information, peculiar characteristics of the tide gauge site (for example, complex local geography, seiching, silting up of the harbour, river mouths) and listings of gaps in the data record. Any problems not resolved with data originators are also documented. Finally the data and accompanying documentation are loaded to the BODC national oceanographic data bank.

In-situ sea level data acquisition and quality control for the 'fast delivery' data essentially parallel those of the 'delayed mode' data. The UHSLC is involved at each level of the data flow from station operation and data acquisition, to quality control and data distribution. The quality assurance methodology was directly influenced by the lessons learned running the UH Sea Level Network. It mirrors the station network data processing steps, and in many cases uses variants of the same software. Some of the concepts incorporated include an emphasis on the use of redundant data and on the timeliness of data acquisition, and the production of 'fully corrected' data. These three issues warrant some discussion.

The 'fast delivery' DAC has aggressively sought redundant data sources. It is not uncommon to find other agencies, such as the Pacific Tsunami Warning Center, who, for their own purposes, have instruments installed at sea level stations. Data from these agencies can serve as important secondary and tertiary sources of information, allowing the DAC to fill gaps and recognize and correct level shifts in the primary originator's data. Additionally, every effort has been made to acquire the data contributed by other agencies as rapidly as possible, often processing satellite-transmitted data as they become available. This has enabled the DAC to give the data originators' agency timely feedback on the quality of their data, thus allowing them to provide additional information that can be used to correct the data. Finally, the UHSLC experience using and distributing the NORPAX data, and subsequently the TOGA data set, has convinced them of the need to construct a 'fully corrected' data set. The vast majority of the investigators which the UHSLC has worked with prefer data that they can use immediately. In fact, to date they have never once had an investigator request the original data set as collected by the UH or provided by an external contributing agency. To this end, erroneous or questionable data are not only flagged, but every effort is made to correct them, and produce a data set that can utilized in projects without any additional processing steps.

Data distribution

Both sea level DACs make their data holdings available via the World Wide Web (WWW), linking into the WOCE Data Management System, and by ftp. A catalogue of data holdings is available from the WOCE Data Information Unit, with links to the two sea level DACs. The UHSLC provides access to the WOCE 'fast delivery' data either via a list of sites or a clickable map - one chooses a site and then has the opportunity to download the data, data documentation and summary of quality assurance procedures. In addition, data from two of the UHSLC's other databases, the Research Quality Database (RQD) and the ISLP (Pacific) database, are also available. For WOCE, hourly values, daily and monthly means are available; for the RQD daily and monthly values, and also data documentation, are available.

'Delayed mode' sea level data are also accessible via the WWW. Again access is available either from a list of sites (one per ocean), via a clickable map, or data from an entire ocean. Selections can be made and either 'zipped' or 'tarred' to allow quicker downloading of the data. Data documentation is also available together with a brief description of procedures. Data prior to 1985 are not yet included on the WWW, these can be requested from BODC. At present hourly data are available; daily means will be added soon. Monthly and annual means can be obtained from the PSMSL, and a link to their data holdings and WWW pages is included.

The UHSLC, together with the US NODC, has produced a CD-ROM of their data holdings of TOGA sea level data; also included is a copy of the PSMSL data set. At the end of the WOCE experiment, BODC will provide WDC-A with a copy of the complete WOCE Sea Level Data Set on CD-ROM but, prior to that, a preliminary version of the data available will be produced at the end of 1997. This CD-ROM will contain the quality controlled data set as held by BODC, most probably in the form described below, as Level 1 and 2 data sets, and will also include the current holdings of the 'fast delivery' centre. In addition, the PSMSL data set and the latest version of the GLOSS Station Handbook will be included. This product will be available to participants in the WOCE Scientific Conference scheduled to take place in Halifax, Canada, in May 1998. Other WOCE DACs will be producing CD-ROMs of their data.

Current sea level DAC status

Data from both WOCE DACs are in high demand. The UHSLC is presently collecting data from 109 stations for the 'fast delivery' component (Figure 1). ORSTOM and Service ARGOS contribute data from 5 stations, the Flinders University in Australia provides data from 12 stations, Canada from 5 stations, the BODC from 6 stations, the U.S. National Ocean Service provides 36 additional stations, and, building upon their contribution to the IGOSS maps, the Japanese Meteorological Agency supplies 8 stations along the coast of Japan. For 66 'fast delivery' stations, the existing time series have been extended backward to 1985 in order to connect GEOSAT and TOPEX/POSEIDON data sets. The 'fast delivery' holdings now include over 1800 site years of hourly data. During the fiscal year 1997, there have been 871 requests for 6527 station data files.

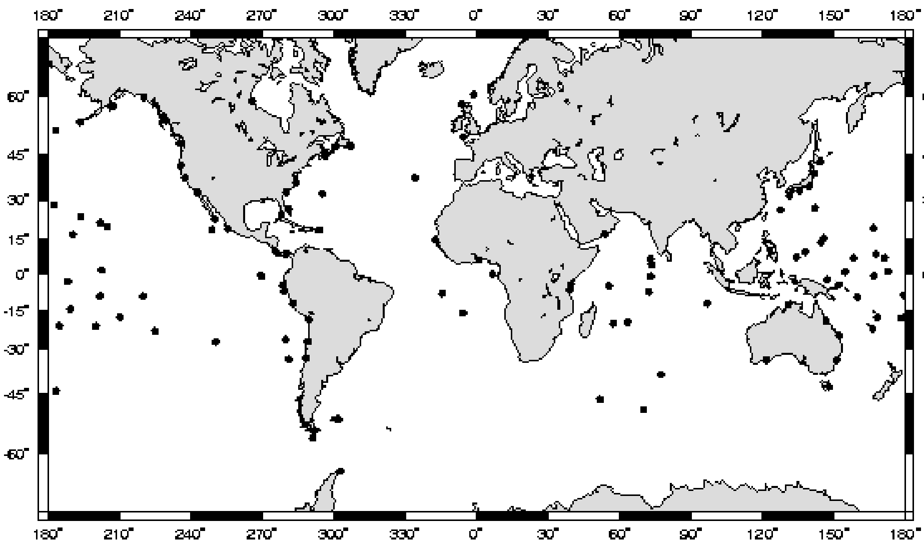

Figure 1. WOCE 'Fast Delivery' Sea Level Stations

The BODC is at present collating data from 156 tide gauge sites worldwide for the WOCE 'delayed-mode' DAC (Figure 2). Data from these have been supplied from 20 countries. The total volume of data received so far is over 2000 site years of which around 900 are post-1984. Data prior to 1985 are held at BODC; in the future these will be added to the Web/ftp site. New additions to the data set have recently been received from Nuuk (Greenland) and Torshavn (Faroes). Additionally the keying of historical manuscript data from Bluff Harbour (New Zealand) and Gibraltar has been completed, although data are continuing to arrive regularly from the Gibraltar gauge. Included in the data set are data from the ACCLAIM network in the South Atlantic (seven island and Antarctic tide gauges plus bottom pressure gauge data from deployments at about 20 locations; so far 45 records, each usually of approximately 1 year in length, have been collected). Work is now underway to acquire the remaining bottom pressure records, and also any inverted echo sounder data.

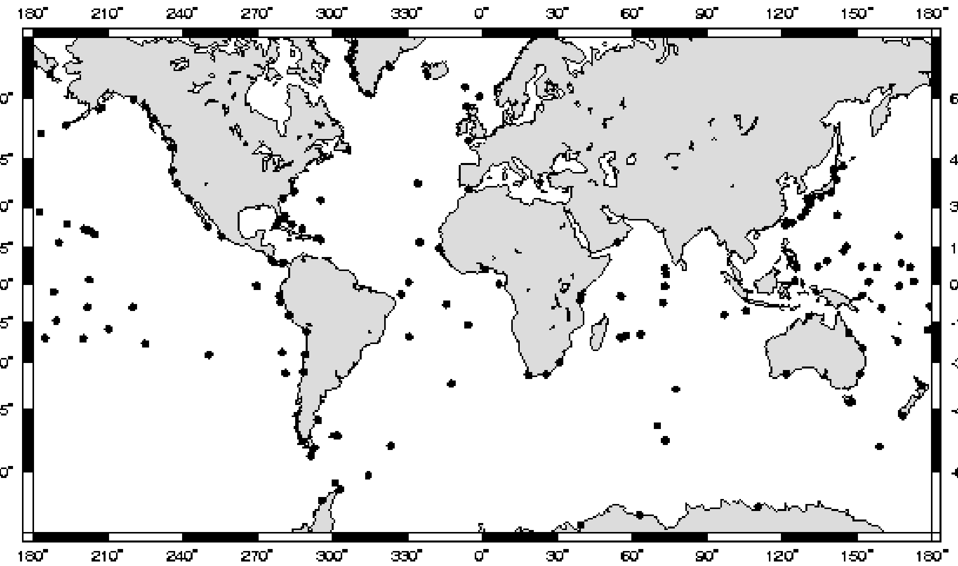

Figure 2. WOCE 'Delayed Mode' Sea Level Stations

Regular requests for sea level data from WOCE scientists are received. However, since the establishment of the public directory, and the Web site, most requests are serviced by the scientists themselves, and a record of them logged into a database. The number of requests over the last year has included 80 logins to the public directory and the retrieval of 550 data files. In addition, the Web site has recorded retrieval of 4353 data files from 188 unique ids since March 1996.

The future for *in-situ* sea level

The Global Sea Level Observing System (GLOSS) is an international system, coordinated by the Intergovernmental Oceanographic Commission (IOC), to provide high quality sea level data from a global network of sea level stations and has strong links with the WOCE sea level centres. The first GLOSS Implementation Plan [10] sought to establish a network of approximately 300 tide gauge stations distributed along continental coastlines and throughout each of the world's island groups. This Implementation Plan has recently undergone a major revision and was approved at the recent IOC Assembly; it will be published soon [11].

One of the main functions of GLOSS is to provide for the smooth flow of sea level data from gauges, to national and international centres, and to scientists. Included in the 1990 Plan was the obligation that countries (or gauges) which were 'committed to GLOSS' should supply monthly and annual values of mean sea level to the Permanent Service for Mean Sea Level (PSMSL), and also make hourly values of sea level data available for international exchange. In the revised Plan, it is strongly requested that countries 'committed to GLOSS' make their hourly or higher frequency data available to one or more of a small number of 'International Sea Level Centres' recognized by GLOSS, or to provide these data sets on an ftp or WWW server in their own organization, in addition to continuing the provision of data to the PSMSL. The higher frequency data are needed to provide the possibility for essential quality control checking of monthly and annual mean sea level values by the PSMSL and to provide access to the higher frequency of the sea level variability spectrum, thereby aiding interpretation of interesting signals which may be less evident in the monthly means, and to enabling long term archiving of these irreplaceable data sets.

The main aim of the above is to make the higher frequency data available with a maximum delay of six months. These data sets may be quality controlled or not, since it is recognized that fully quality controlling the data sets can take some time. The data sets will be marked accordingly. Each Centre will keep detailed log files documenting data sources, quality control histories, and changes and revisions to data sets.

In addition, there will be an on-going need for the type of operation undertaken by the UHSLC for WOCE to support further altimetry work. This Centre will need access to the higher frequency tide gauge data within a month or so of collection. The GLOSS network contains a recommended subset of gauges for altimetry work which, for most gauges, should be able to meet these requirements.

Each International Sea Level Centre recognised by GLOSS will be required to provide full on-line catalogues of both their own data holdings and those of the other Centres. In addition, they must provide ftp or WWW addresses for all national authorities with their own servers. Consequently, if requested data are not available from one site, they can be immediately accessed over the Internet from another.

A GLOSS Data Coordination Panel is being set up which has the responsibility to ensure that this coordination takes place. This Panel, building on the WOCE experiences of the UHSLC and BODC, will establish standards (for quality control and data formats) by which sea level centres, either international or national, can present their GLOSS sea level data in a common way to users. Taking into account the modes of operation of the two WOCE sea level centres, the likely way forward will now be described.

It is suggested that sea level data are made available at several 'levels' and, borrowing the terminology used in remote sensing, sea level products will be defined as Level 0, 1, 2, etc., as follows:

Level 0 data set

Raw data as received from the gauges. These are data sets being worked on, perhaps being calibrated, and of use primarily only to the data collector. They would obviously be in the local format. At this level there is little coordination between agencies and centres, and these data are not made available.

Level 1 data set

This data set contains files which have physical units, and have undergone quality control. In BODC terms they will have been flagged to indicate null, suspect, or interpolated values, and also datum changes. But no data values will have been changed without the permission of the data collector. A Level 1 data set preserves the original data in terms of the sampling interval (e.g. 6 minute, 15 minute, etc.), and parameters measured (e.g. any ancillary measurements, temperature, wind speed/direction, atmospheric pressure). These data will eventually be copied to an internationally-agreed archiving centre for safety, and will be available for scientific users (for example to study seiches or tsunamis).

Using Newlyn, UK, as an example, Table 2 illustrates the contents of a Level 1 data set. Within the sea level data files, data will be arranged in whatever format is optimum for the institution as long as it is defined clearly (units, time zones, datums) in the corresponding document file. These document files must contain full datum information and a quality control history. It is also recommended that location maps and photographs of the site be included. If limited (or no) quality control has been carried out, the documentation should indicate this.

Level 2 data sets

The most usual Level 2 data set would be fully quality controlled and commonly-formatted tide gauge data. Included in a Level 2 data set would be the following:

(1) hourly values of sea level, centred on the hour and in UT. These data should form a continuous time series with respect to a land benchmark datum.

(2) daily and monthly mean sea level values.

(3) documentation files, including a quality control history table, together with a derivation history (i.e. from which Level 1 files they were derived).

The data will be fully quality controlled using guidelines now under development based on the experiences of the WOCE DACs and the recommended format is that used by the UHSLC (Table 3), although small amendments may be made to this. For example, the 3 plus 8 bytes in the data lines of the hourly files for UH station number and station name could be generalised into an 11 byte station code or name as preferred by the centre. The documentation will also follow a recommended structure.

Level 3 data sets

These have to be further considered in the future, but an obvious one could be the 'revised local reference' (RLR) type sea level (as produced by the PSMSL) corrected for land movements using GPS information, as discussed at the joint IGS/IOC/PSMSL workshop on geodetic fixing of tide gauge benchmarks (March 1997)[12].

The new GLOSS Implementation Plan suggests one or more international archiving centres which will copy the files from each of the international and national centres at regular, agreed intervals, most probably annually - but also when a centre ends its operation. This data set will form the GLOSS Archive. The archiving centres can merge the inventory information, according to a pre-determined set of rules, for example to remove duplicate entries. This inventory will be made available via the WWW and ftp, pointing to the international and national sites where the data are available. In time the GLOSS Archive can also be made available as a whole.

Conclusion

The two WOCE sea level DACs are working well and both have built up expertise in the collection, quality assurance and management of sea level data. Slight differences in working practices at the two centres has resulted in improvements in procedures at both centres, and an increase in collaboration between them. Large volumes of data are readily available at both centres, and CD-ROM containing these data will be produced at the end of 1997.

Together, and in collaboration with the GLOSS programme, the UHSLC and BODC are developing standards for sea level data (e.g. guidelines for quality control, data documentation and exchange formats). Following the guidelines of the GLOSS implementation Plan, it is possible that in the future, the 'fast delivery' centre will evolve into the suggested GLOSS altimetry centre and the 'delayed-mode' centre, together with PSMSL, will become a GLOSS Archive site.

Acknowledgements

The authors would like to record our appreciation of all the technicians, tide gauge operators, digitizers, data processors and coordinators who have helped to make the sea level data available to the DACs, and especially we would like to thank our colleagues at the two DACs who handle the data on a day to day basis.

References

1. Shaw, S.M. 1993. Class A Network Dataring Gauges 1992 data processing and analysis. Proudman Oceanographic Laboratory Report No 30, 178pp.

2. Rickards, L.J. The MEDALPEX Sea Level Data Set. Annales Geophysicae, 5B(1), 13-20.

3. Wyrtki, K., and P. Caldwell, K. Constantine, B. J. Kilonsky, G. T. Mitchum, B. Miyamoto, T. Murphy, and S. Nakahara, 1988, The Pacific island sea level network. Univ. of Hawaii, JIMA 88-0137, 71pp.

4. Kilonsky, B and P. Caldwell, 1991, In the pursuit of high-quality sea level data. in IEEE Oceans Proceedings, 2, 669-675.

5. Wyrtki, K., B. Kilonsky, and S. Nakahara, 1988, The IGOSS pilot project in the Pacific. Univ. of Hawaii, JIMAR 88-0150.

6a. World Ocean Circulation Experiment (WOCE). 1988. World Ocean Circulation Experiment Implementation Plan. Volume 1: Detailed requirements. WMO World Climate Research Programme, WCRP-11, 63pp. (WMO/TD-No.242).

6b. World Ocean Circulation Experiment (WOCE). 1988. World Ocean Circulation Experiment Implementation Plan. Volume 2: Scientific Background. WMO World Climate Research Programme, WCRP-12, 130pp. (WMO/TD-No.243).

7. WOCE Data Information Unit. 1994. WOCE Data Handbook (second printing) Looseleaf. WOCE International Project Office. WOCE Report 120/94. 180pp

8. Mitchum, G., B. Kilonsky and B. Miyamoto, 1994, Methods for maintaining a stable datum in a sea level monitoring system. In IEEE Oceans Proceedings, September, I25-30.

9. Kilonsky, B., 1984, Satellite-transmitted sea level stations - 1983. Ref. HIG-84-2, Univ. Of Hawaii, Honolulu, 40 pp.

10. IOC. 1990. Global Sea Level Observing System (GLOSS) Implementation Plan. Intergovernmental Oceanographic Commission, Technical Series, No. 35, 90pp.

11. IOC. 1997. Global Sea Level Observing System (GLOSS) Implementation Plan. Intergovernmental Oceanographic Commission, Technical Series. (in press)

12. IGS/IOC/PSMSL 1997. Report of the IGS/IOC/PSMSL workshop on geodetic fixing of tide gauge benchmarks. (in preparation)

Figure and table captions

Figure 1. WOCE 'fast delivery' sea level data stations

Figure 2. WOCE 'delayed-mode' sea level data stations

Table 1. Information required to accompany *in situ* sea level data

Table 2. Example of Level 1 directory for Newlyn data from 1985 to 1996

Table 3. University of Hawaii Sea Level Center hourly sea level data format

Table 1. Information required to accompany *in situ* sea level data

Each sea level data series should be accompanied by the following information:

Country and organisation responsible for data collection and processing

Originator's identifier for the series (e.g. site name and year)

Geographical location (latitude and longitude)

Dates and times of the start and end of the data series

Sufficient plain language documentation should accompany the data so as to ensure that they are adequately qualified and may therefore be used with confidence by a secondary user.

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.

Site

Brief description of location of tide gauge

Description of tide gauge benchmarks

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

Report any additional item or event that may have affected the data, or have a bearing on the subsequent use of the data.

Table 2. Example of Level 1 directory for Newlyn data from 1985 to 1996

The Level 1 directory contains the following files:

INVENTORY1: contains a set of 'masterfile' names relating to tide gauge data (in the most general sense), as follows:

station latitude longitude masterfilename1 start end source GLOSS No.

station latitude longitude masterfilename2 start end source GLOSS No.

Thus, an INVENTORY1 file for Newlyn data from 1985-1996 looks as follows:

Station	Latitude dd mm.mh	Longitude ddd mm.mh	Masterfile name	Start Date dd-Mon-yyyy	End Date dd-Mon-yyyy	Source	GLOSS No.
Newlyn	50 06.1N	005 32.5W	new1985.inv	01-Jan-1985	31-Dec-1985	POL	241
Newlyn	50 06.1N	005 32.5W	new1986.inv	01-Jan-1986	31-Dec-1986	POL	241
Newlyn	50 06.1N	005 32.5W	new1987.inv	01-Jan-1987	31-Dec-1987	POL	241
Newlyn	50 06.1N	005 32.5W	new1988.inv	01-Jan-1988	31-Dec-1988	POL	241
Newlyn	50 06.1N	005 32.5W	new1989.inv	01-Jan-1989	31-Dec-1989	POL	241
Newlyn	50 06.1N	005 32.5W	new1990.inv	01-Jan-1990	31-Dec-1990	POL	241
Newlyn	50 06.1N	005 32.5W	new1991.inv	01-Jan-1991	31-Dec-1991	POL	241
Newlyn	50 06.1N	005 32.5W	new1992.inv	01-Jan-1992	31-Dec-1992	POL	241
Newlyn	50 06.1N	005 32.5W	new1993.inv	01-Jan-1993	31-Dec-1993	POL	241
Newlyn	50 06.1N	005 32.5W	new1994.inv	01-Jan-1994	31-Dec-1994	POL	241
Newlyn	50 06.1N	005 32.5W	new1995.inv	01-Jan-1995	31-Dec-1995	POL	241
Newlyn	50 06.1N	005 32.5W	new1996.inv	01-Jan-1996	31-Dec-1996	POL	241

masterfilename.inv contains a list of data type identifiers and filenames.

Each of the files in the masterfile is automatically defined as being relevant to the start/end periods of the inventory. The types of files to be found in the masterfilename.inv are listed below. This list also serves as a note of the other files to be found in the directory.

filename.slr	contains 'raw' (i.e. high frequency) sea level data
filename.ssp	sub-surface pressure data
filename.sap	atmospheric pressure data
filename.st	sea temperature data
filename.at	air temperature data
filename.gps	GPS data relevant to this station
filename_lev.doc	levelling data relevant to this station
filename_slr.doc	site history documentation file
filename_img.tif	image file(s) referred to in filename_slr.doc and dated (eg location maps, photographs)

Contents of Level 1 directory for Newlyn, UK (1985-1996) is as follows:

INVENTORY1	new1985.inv	new1991.inv	new1985.slr	new1991.slr
	new_img1.tif	new1986.inv	new1992.inv	new1986.slr
	new1992.slr	new_img2.tif	new1987.inv	new1993.inv
	new1987.slr	new1993.slr	new_slr.doc	new1988.inv
	new1994.inv	new1988.slr	new1994.slr	new_lev.doc
	new1989.inv	new1995.inv	new1989.slr	new1995.slr
	new1990.inv	new1996.inv	new1990.slr	new1996.slr

Contents of a .inv file, for example, new1989.inv

new1989.slr
new_lev.doc
new_slr.doc
new_img1.tif
new_img2.tif

Table 3. University of Hawaii Sea Level Centre hourly sea level data format

Each file is given a name "h###.dat" where "h" denotes hourly sea level data and "###" denotes the station number. A file exists for every station with hourly data.

Within the file, each year begins with a header record. The header includes the station number (columns 1-3), the station name (columns 4-11), year (columns 12-15), latitude (columns 22-25), longitude (columns 37-45) and time zone (columns 57-60).

This header is followed by the data records, each containing 12 hourly values, half a day of data. These records contain the station-id (columns 1-11), year (columns 12-15), month (columns 16-17), day (columns 18-19), record count, either 1 for hours 0:00 - 11:00, or 2 for hours 12:00 - 23:00 (columns 20), and 12 hourly sea level values (1st value columns 21-25, 2nd value columns 26-30,....., and 12th value columns 76-80).

Data are in units of millimetres, missing data values are given a value of 9999, and datum information is provided in the accompanying documentation file.

An example is given below for: **NUKU HIVA, French Polynesia** (UHSLC station number 031)

```
031NUKU HIV1996  LAT=08 55.8S  LONG=140 04.9W  TIMEZONE=GMT
031NUKU HIV1996  1  11 1698 1453 1236 1102 1004 1070 1289 1520 1800 1985 2073 2044
031NUKU HIV1996  1  12 1876 1654 1442 1211 1081 1095 1221 1436 1648 1831 1935 1965
031NUKU HIV1996  1  21 1852 1677 1417 1229 1074 1046 1176 1405 1647 1893 2051 2129
031NUKU HIV1996  1  22 2045 1879 1648 1410 1199 1090 1132 1290 1523 1745 1897 1990
031NUKU HIV1996  1  31 1954 1798 1591 1353 1162 1044 1045 1186 1403 1663 1894 2061
031NUKU HIV1996  1  32 2084 1991 1772 1526 1267 1114 1093 1138 1298 1535 1775 1916
031NUKU HIV1996  1  41 1965 1894 1737 1501 1284 1111 1017 1076 1282 1510 1788 2033
031NUKU HIV1996  1  42 2116 2097 1942 1758 1500 1252 1103 1078 1215 1419 1631 1805
031NUKU HIV1996  1  51 1957 1978 1879 1685 1455 1241 1083 1044 1141 1355 1649 1904
031NUKU HIV1996  1  52 2102 2177 2101 1913 1653 1387 1202 1048 1117 1316 1526 1743
031NUKU HIV1996  1  61 1902 1979 1945 1799 1627 1363 1140 1041 1098 1253 1486 1722
031NUKU HIV1996  1  62 1950 2073 2099 1997 1787 1530 1309 1126 1081 1162 1317 1577
031NUKU HIV1996  1  71 1790 1913 1932 1894 1737 1500 1298 1128 1063 1127 1326 1530
031NUKU HIV1996  1  72 1798 2019 2092 2038 1908 1678 1434 1230 1118 1132 1256 1466
031NUKU HIV1996  1  81 1677 1862 1939 1932 1851 1663 1452 1255 1125 1095 1218 1412
031NUKU HIV1996  1  82 1647 1897 2012 2073 2016 1850 1615 1386 1204 1137 1157 1306
031NUKU HIV1996  1  91 1507 1707 1843 1917 1914 1772 1609 1347 1201 1119 1132 1282
031NUKU HIV1996  1  92 1507 1746 1971 2099 2112 1983 1788 1532 1305 1159 1151 1231
031NUKU HIV1996  1101 1385 1584 1810 1947 1967 1895 1743 1545 1355 1205 1169 1243
031NUKU HIV1996  1102 1374 1592 1799 1985 2075 2028 1891 1676 1468 1272 1168 1161
031NUKU HIV1996  1111 1287 1462 1642 1827 1939 1943 1886 1732 1527 1362 1251 1209
031NUKU HIV1996  1112 1267 1455 1663 1861 2011 2080 1991 1860 1616 1433 1267 1170
031NUKU HIV1996  1121 1178 1318 1484 1640 1817 1889 1907 1817 1665 1500 1321 1189
031NUKU HIV1996  1122 1158 1267 1456 1661 1862 1967 1980 1909 1760 1586 1359 1186
031NUKU HIV1996  1131 1121 1164 1298 1480 1676 1826 1912 1891 1804 1613 1453 1298
031NUKU HIV1996  1132 1232 1256 1318 1486 1708 1881 1977 2003 1924 1760 1545 1354
031NUKU HIV1996  1141 1216 1146 1190 1343 1536 1747 1889 1991 1967 1897 1727 1529
031NUKU HIV1996  1142 1377 1267 1275 1356 1532 1746 1884 1987 1990 1893 1729 1517
031NUKU HIV1996  1151 1307 1175 1111 1180 1367 1571 1797 1974 2025 1993 1885 1696
031NUKU HIV1996  1152 1487 1283 1168 1189 1305 1467 1679 1859 2012 2020 1907 1740
```