The challenge of providing metadata for a 200 year long global mean sea level dataset

Andrew Matthews, Permanent Service for Mean Sea Level, National Oceanography Centre (PSMSL-NOC) (United Kingdom), antt@noc.ac.uk
Lesley Rickards, PSMSL-NOC / British Oceanographic Data Centre, NOC (United Kingdom), ljr@bodc.ac.uk
Elizabeth Bradshaw, British Oceanographic Data Centre, NOC (United Kingdom), elizb@bodc.ac.uk
Kathy Gordon, PSMSL-NOC (United Kingdom), kmg@noc.ac.uk
Angela Hibbert, PSMSL-NOC (United Kingdom), anhi@noc.ac.uk
Svetlana Jevrejeva, PSMSL-NOC (United Kingdom), sveta@noc.ac.uk
Simon Williams, PSMSL-NOC (United Kingdom), sdwil@noc.ac.uk
Philip Woodworth, PSMSL-NOC (United Kingdom), plw@noc.ac.uk

The Permanent Service for Mean Sea Level Dataset

The Permanent Service for Mean Sea Level (PSMSL) provides an internationally recognised focus for knowledge and expertise on sea level science, and constructs and maintains a global data set of long term sea level change information from tide gauges. The PSMSL data bank holds over 67,500 station-years of monthly and annual mean sea level data from over 2300 tide gauge stations, some dating back to the start of the 19th century. Data from each site are quality controlled and, wherever possible, reduced to a common datum, whose stability is monitored through a network of geodetic benchmarks. All data are freely available from the PSMSL website, and are used in a wide range of disciplines, including oceanography, geology, geodesy and climate change studies. The spatial distribution and length of each series in the PSMSL dataset is shown in Figure 1.

The PSMSL was established as a "Permanent Service" of the International Council of Scientific Unions (ICSU) in 1958, but in practice was a continuation of the Mean Sea Level Committee which had been set up at the Lisbon International Union of Geodesy and Geophysics (IUGG) conference in 1933. Today, the PSMSL operates under the auspices of the International Council for Science (the current name of ICSU), and is a regular member of the ICSU World Data System.

Metadata Challenges

The sea level record at a location is the result of many years of effort. Over this period many changes in the method of observation will have been made. Several sensors of different types may have been used, perhaps at different locations, and the organisation responsible for taking the measurements may have changed on one or more occasions.

Throughout the observational period the stability of the tide gauge site must be monitored to prevent artificial trends being introduced into the series, usually through annual geodetic levelling. Additionally, larger scale land motion can result from factors including earthquakes, volcanic activity, glacial isostatic adjustment and groundwater extraction.

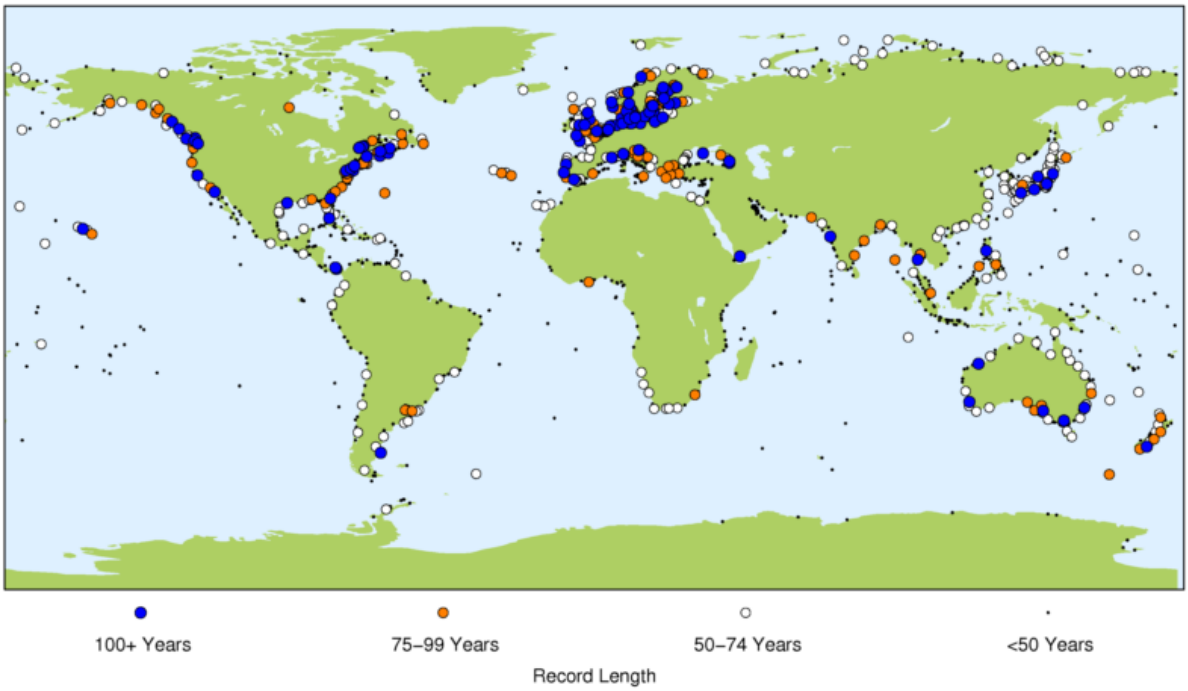

Figure 1: Length of Mean Sea Level records in the PSMSL database

Before including the data in the PSMSL dataset, the high frequency data must be reduced to monthly mean data in a way that removes any tidal cycle. The sampling rate of high frequency data may have changed, and in some older series observations may only have been made in daylight hours.

As a result, providing a coherent set of metadata that summarises the entire history of sea level measurement at a site, including all the caveats that users of our data should be aware of, can be highly challenging, particularly when historical records can be incomplete, vague, or even contradictory.

As may be expected from an organisation with an over 80 year history, change occurs slowly, and as a result PSMSL has only recently begun to make efforts to provide metadata meeting internationally approved standards. Currently, large amounts of metadata are still presented as blocks of plain text rather than structured information. Other information is not yet publically accessible, in some cases only existing in PSMSL's extensive paper archive.

Here we will describe the PSMSL dataset and our efforts to provide well structured and standardized metadata. Data will soon be distributed in CF-compliant netCDF format. Furthermore, we have been investigating the feasibility of using various OGC standards, including Observations and Measurements (O&M) and SensorML. We will illustrate some of the challenges faced in trying to adopt standards designed to describe 21st century technologies to represent data collected using methods that have been evolving since the 19th century.

This presentation will not offer a complete solution to these problems, merely some steps towards the latest evolution of an ongoing endeavour to accurately describe the changing height of the ocean. Whatever solutions are adopted will need to meet the requirements of the International Association of Geodesy's Global Geodetic Observing System (GGOS) and the Intergovernmental Oceanographic Commission's Global Sea Level Observing System (GLOSS), and also the needs of other users of our data.