Using citizen science to rescue tide gauge data

Andrew Matthews, National Oceanography Centre (United Kingdom), antt@noc.ac.uk
Elizabeth Bradshaw, National Oceanography Centre (United Kingdom), elizb@bodc.ac.uk
Kathy Gordon, National Oceanography Centre (United Kingdom), kmg@noc.ac.uk

The need for tide gauge data rescue

Tide gauges are the main source of data that underpin our knowledge of coastal sea level. They have been essential to our study of tidal science, but also how sea level means and extremes will change as sea level rises. They make a vital contribution to the study of ocean currents and local land movement, and also navigation around harbours (Marcos et al., 2019). Unfortunately, the distribution of stations in the global tide gauge network is heavily weighted towards the Europe and North America, particularly when considering long term installations (Figure 1).

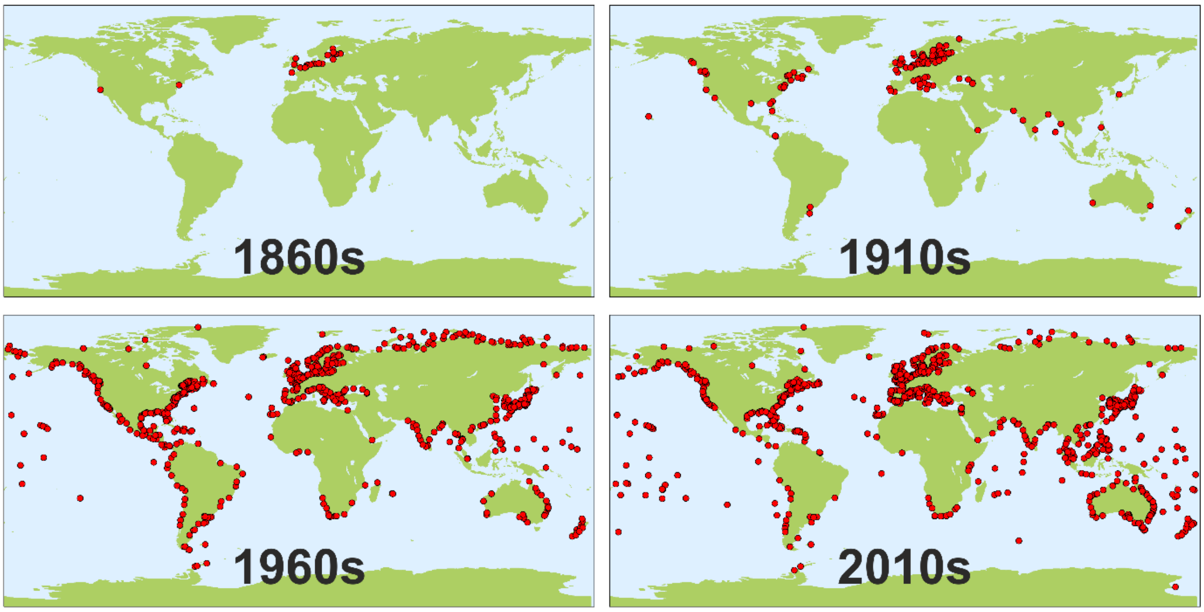

Figure 1: The development of the global tide gauge network over time, as held by PSMSL (https://psmsl.org)

There is a clear need for extra data to improve our knowledge in data sparse areas, but also to improve the quality of our predictions by extending the length of the time series available. Such data already exists in the form of charts and ledgers that have never been converted into a digital format (Figure 2). The International Oceanographic Commission's Global Sea Level Observing System (GLOSS) programme has been encouraging organistations to search for such material, and investigating efficient methods of digitisation (Bradshaw et al., 2015).

Digitisation Difficulties

There have recently been several successful studies extending long records via the digitisation of old data, for example, Talke et. al (2018) digitised 50 years of historical data to extend the record at Boston, USA to cover the period 1825-2018. However, the amount of effort can be significant: that example required the recovery of data and metadata from several national, state, and university libraries and archives, and a team of student digitisers to process them. The use of automated scanning using optical character recognition enchanced by machine learning seems an obvious solution, however current attempts have been thwarted by regular changes in formats of historical material, and by the variations of handwritten characters and numerals, particularly in older documents.

Citizen Science : a possible solution ?

One possible avenue to explore is the use of large groups of volunteers to digitise data, an approach which has been successful in related areas. For example, Hawkins et al. (2019) describes an effort to digitise 1.5 millions of observations taken from the summit of Ben Nevis (the point of maximum elevation in the UK) between 1883-1904: the project was completed in under 3 months by over 1500 volunteers. These projects require careful thought in how to make the project understandable, attractive, and interesting to potential volunteers: in this case the tale of a team of three observers living at the top of Scottish mountain throughout the year undoubtedly helped.

Large online platforms such as https://www.zooniverse.org/ provide the infrastructure required to create these projects and accumulate the volunteers' efforts. The team behind the Ben Nevis project are currently operating a project to digitised monthly rainfall totals from the UK between the 1820s and the 1960s (https://www.zooniverse.org/projects/edh/rainfall-rescue), which at the time of writing has digitised about 4 millions observations in a week with the help of about 15,000 volunteers.

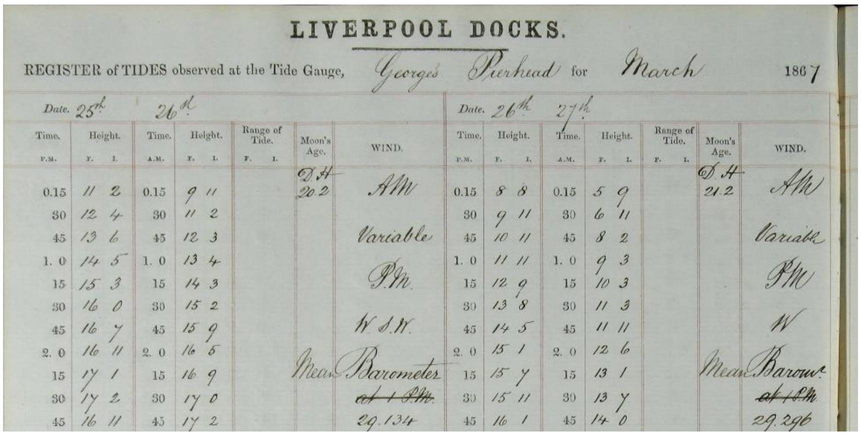

Figure 2 : An example of a ledger containing tide gauge data

We are in the process of creating a citizen science project to digitise sea level data. The project is in currently in active development, and will focus on the recovery of data from approximately 15,000 ledgers recording either 15 minutes observations or the height and time of high and low waters from two sites in the Merseyside area: the historic George's Dock in Liverpool, and Hilbre Island, a tidal island off the Wirral peninsula. We will need to define the process volunteers will undertake, tackle challenges in converting outputs to usable data, and create a plan to archive it in a long term data bank. Possible future extensions of the project include digitising other areas, and including tidal charts in future citizen science projects.

References

Bradshaw, E., Rickards, L. & Aarup, T. (2015). Sea level data archaeology and the Global Sea Level Observing System (GLOSS). GeoResJ, 6, 9-16. https://doi.org/10.1016/j.grj.2015.02.005

Hawkins, E., Burt, S., Brohan, P. et al. (2019). Hourly weather observations from the Scottish Highlands (1883–1904) rescued by volunteer citizen scientists. Geoscience Data Journal, 6 (2), 160-173. https://doi.org/10.1002/gdj3.79

Marcos, M., Wöppelmann, G., Matthews, A. et al. (2019) Coastal Sea Level and Related Fields from Existing Observing Systems. Surv Geophys 40, 1293–1317. https://doi.org/10.1007/s10712-019-09513-3

Talke, S. A., Kemp, A. C., & Woodruff, J. (2018). Relative sea level, tides, and extreme water levels in Boston harbor from 1825 to 2018. Journal of Geophysical Research: Oceans, 123, 3895– 3914. https://doi.org/10.1029/2017JC013645