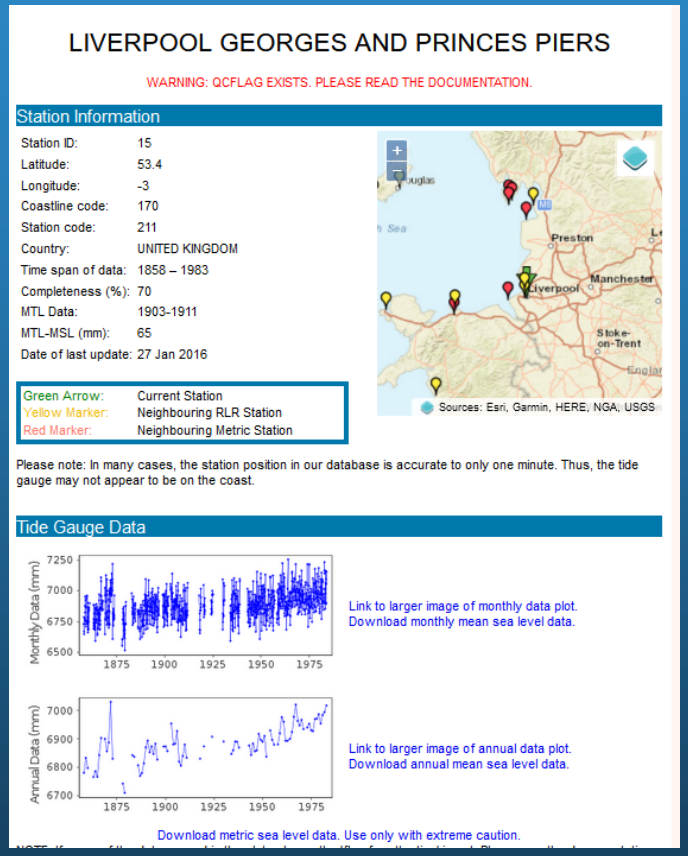
# Using citizen science to digitise 3 million hand-written tide-gauge data entries

**JOANNE WILLIAMS**, ANDREW MATTHEWS, ELIZABETH BRADSHAW & MORE THAN 3831 VOLUNTEERS

National Oceanography Centre, Liverpool, UK. joll@noc.ac.uk



EGU General Assembly
23rd May 2021
EGU22-740
Sharing is encouraged



LIVERPOOL GEORGES AND PRINCES PIERS

WARNING: QCFLAG EXISTS. PLEASE READ THE DOCUMENTATION.

Station Information

| | |
|---|---|
| Station ID: | 15 |
| Latitude: | 53.4 |
| Longitude: | -3 |
| Coastline code: | 170 |
| Station code: | 211 |
| Country: | UNITED KINGDOM |
| Time span of data: | 1858 – 1983 |
| Completeness (%): | 70 |
| MTL Data: | 1903-1911 |
| MTL-MSL (mm): | 65 |
| Date of last update: | 27 Jan 2016 |



| | |
|---|---|
| Green Arrow: | Current Station |
| Yellow Marker: | Neighbouring RLR Station |
| Red Marker: | Neighbouring Metric Station |

Please note: In many cases, the station position in our database is accurate to only one minute. Thus, the tide gauge may not appear to be on the coast.

Tide Gauge Data



Link to larger image of monthly data plot.
Download monthly mean sea level data.

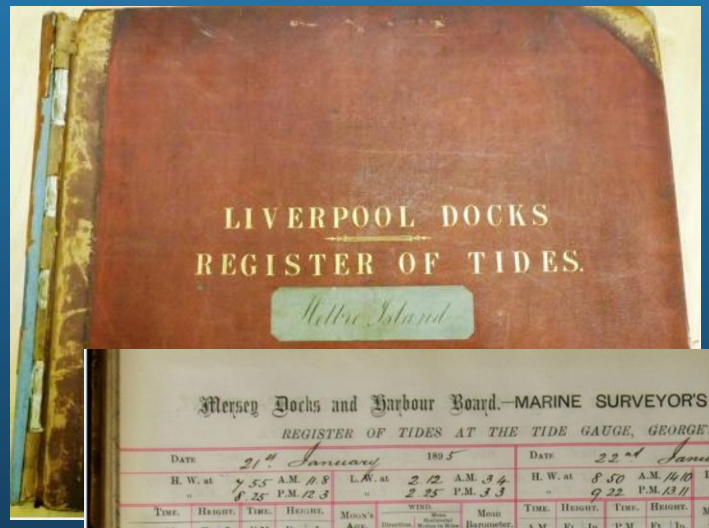


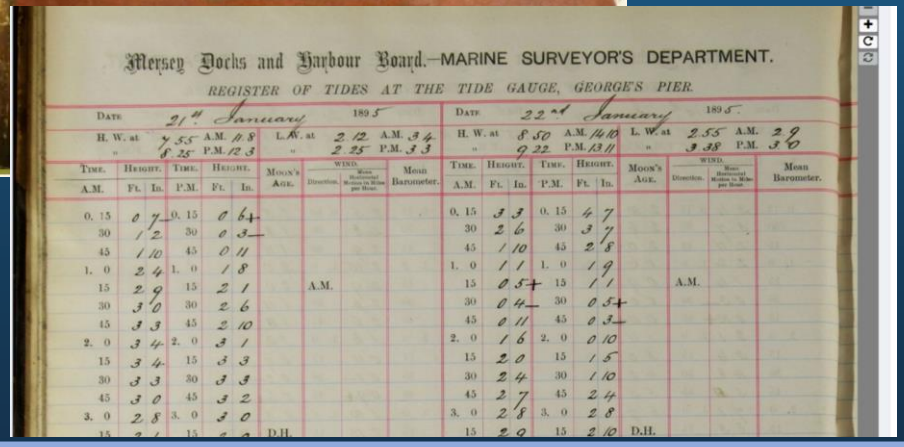
Link to larger image of annual data plot.
Download annual mean sea level data.

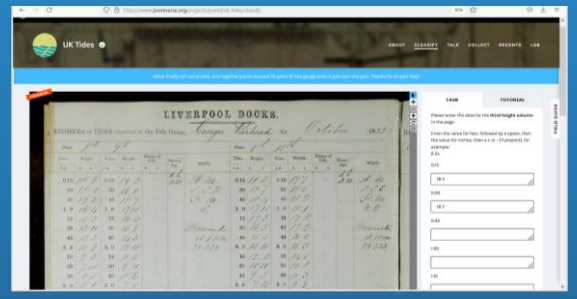
Download metric sea level data. Use only with extreme caution.

Liverpool is one of the longest sea-level records in PSMSL. But the only digital record from the 1800s is hand-calculated monthly mean data, with many gaps.

The high frequency (15 minute) records contain much more information, including tides and storm surges. They still exist, in hand-written ledgers.
We scanned 16000 pages from Hilbre Island and Liverpool, dated 1857 to 1903. It would take years to transcribe!




Fascinating to discover how pivotal this location seems to be.

In 2021, we guided volunteers at zooniverse.org to transcribe the records. We have compared their answers, at least 5 for each entry, and are now working on quality controlling the data.

# Using citizen science to digitise 3 million hand-written tide-gauge data entries

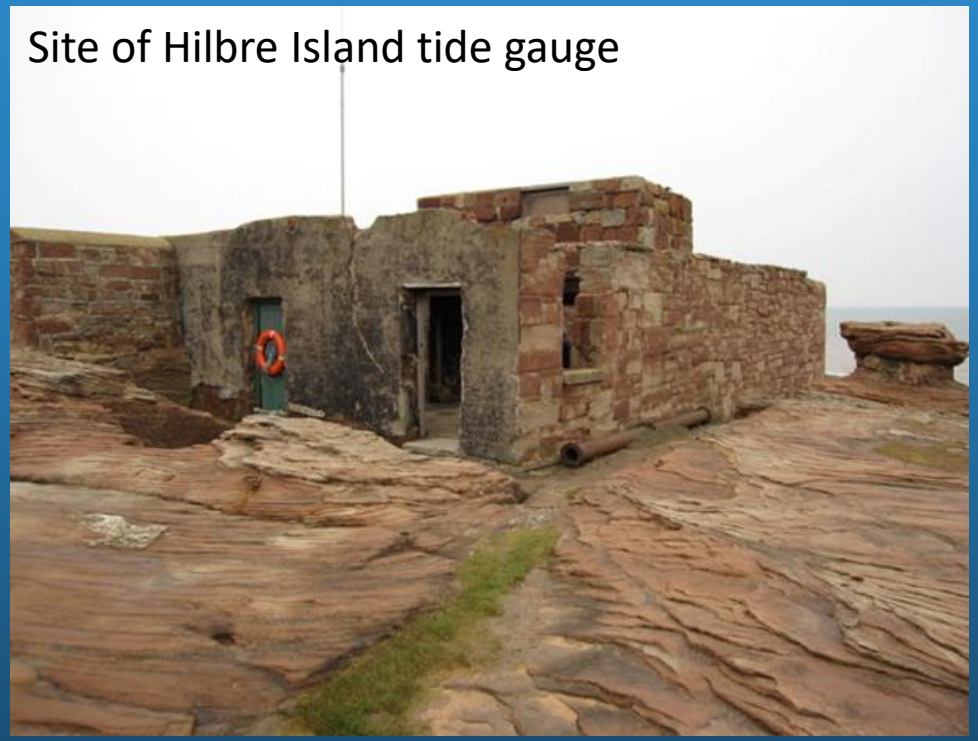
**JOANNE WILLIAMS**, ANDREW MATTHEWS, ELIZABETH BRADSHAW & MORE THAN 3831 VOLUNTEERS

National Oceanography Centre, Liverpool, UK. joll@noc.ac.uk

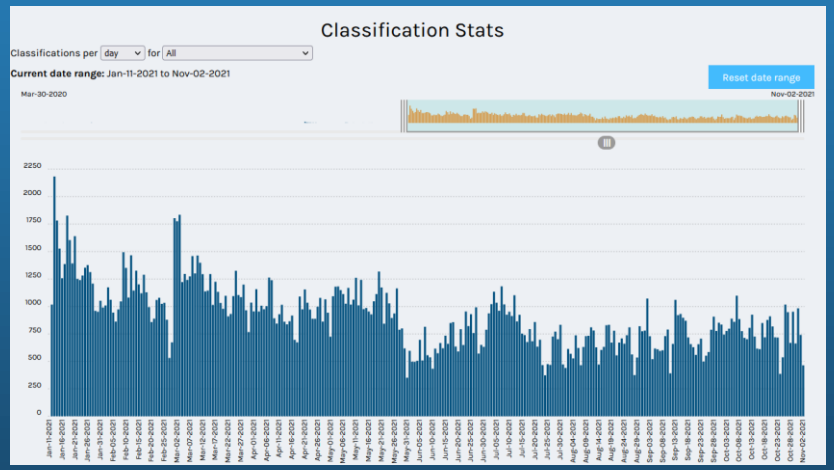


EGU General Assembly
23rd May 2021
EGU22-740

Site of Hilbre Island tide gauge



Zooniverse.org hosts citizen science projects as diverse as identifying galaxies and birds to reading historical letters. It is free to use.
Several volunteers can see each image so you can cross-check entries.
Volunteers can do as much as they like.
New projects are tested before going public




- At over 6000 classifications per week - including Christmas Day - the volunteers worked at the rate of about 5 full time experts.
- It took just over a year to rescue about 100 station-years of 15 minute data.
- Our work has included pre-processing scans, zooniverse setup, publicity, and 2000 responses to volunteer queries.
- We are still working on the quality control and data analysis.
  https://www.zooniverse.org/projects/psmsl/uk-tides

…being a 100% disabled veteran this gives me something to do besides play games and read depressing news stories. I'm more than happy to help.

# Using citizen science to digitise 3 million hand-written tide-gauge data entries

**JOANNE WILLIAMS**, ANDREW MATTHEWS, ELIZABETH BRADSHAW & MORE THAN 3831 VOLUNTEERS

National Oceanography Centre, Liverpool, UK. joll@noc.ac.uk

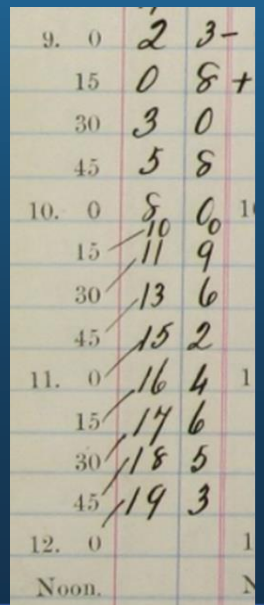


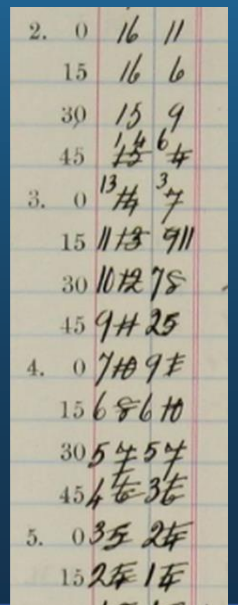
EGU General Assembly
23rd May 2021
EGU22-740

## "Can't you use AI text recognition?"

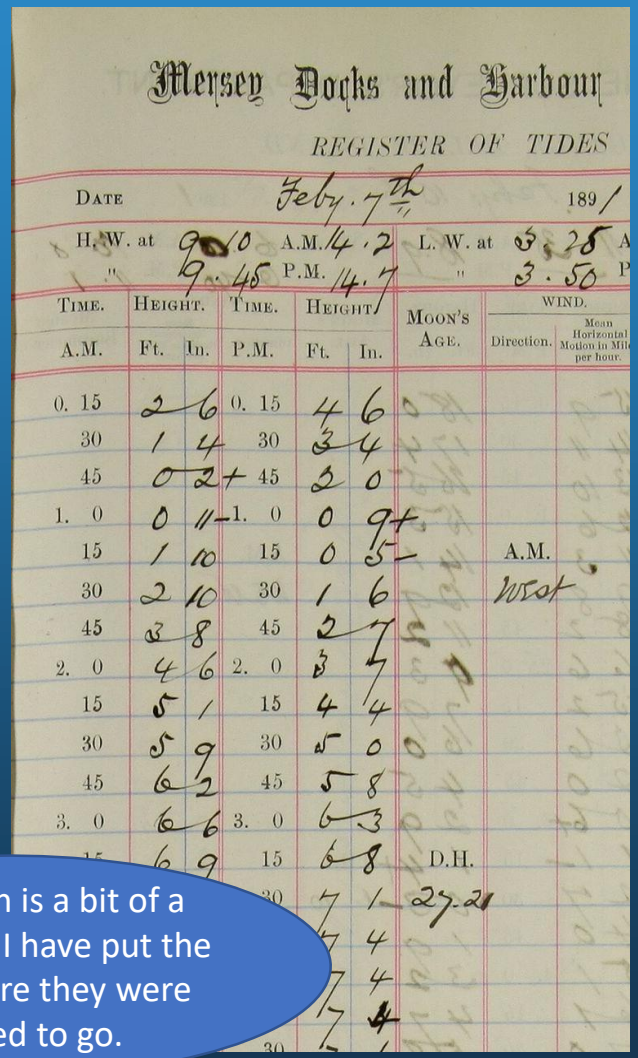
We have seen automatic transcription which is close to capturing the neater pages. But there are many pages which would be very difficult. Some combination of text recognition and volunteer checking is more likely to succeed.

Really terrible handwriting

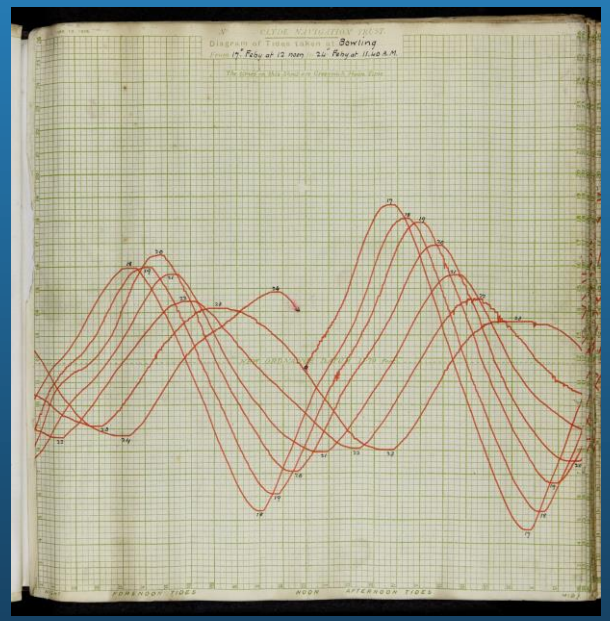




This column is a bit of a mess. I think I have put the figures where they were intended to go.




At some sites, original paper charts exist. They are hard to read even for experts. But if they can be digitised they could yield even higher frequency data.



One challenge is still to locate, catalogue and scan all these records before they are lost.

# Using citizen science to digitise 3 million hand-written tide-gauge data entries

**JOANNE WILLIAMS**, ANDREW MATTHEWS, ELIZABETH BRADSHAW & MORE THAN 3831 VOLUNTEERS

National Oceanography Centre, Liverpool, UK. joll@noc.ac.uk

EGU General Assembly
23rd May 2021
EGU22-740

## Additional Quality Control

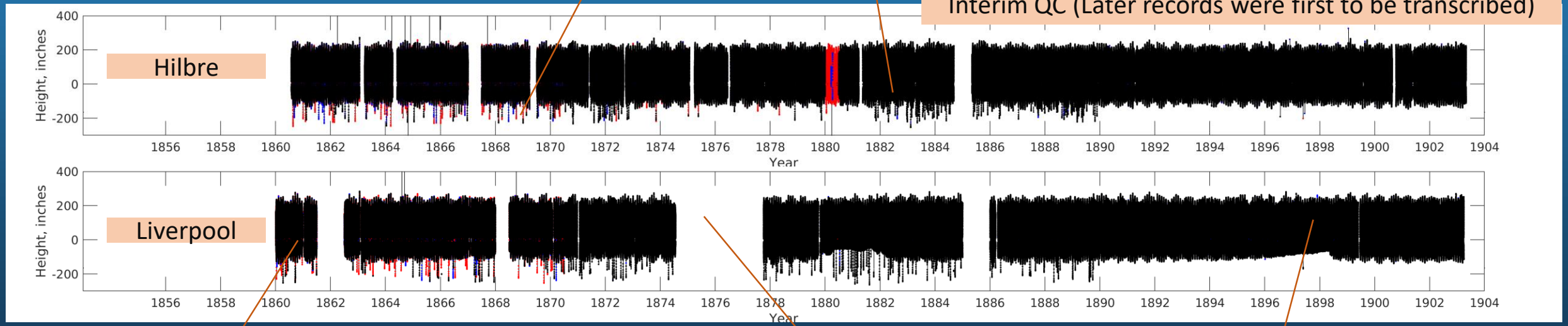
- Assigning correct dates to each page
- Clock drift
- Missing digits (10 feet errors) in ledgers
- Flushing of sediment in pipe at Hilbre
- Hilbre Datum

Disagreement between volunteers awaiting arbitration

Incorrect signs in original record. Can be corrected in QC

UK Tides data entry is a great way to sober up after a couple of pints at the pub

Interim QC (Later records were first to be transcribed)




1857-1869 transcription since completed

Gap at Liverpool due to dock fire. Can sea level here be reconstructed using neighbouring Hilbre?

Siltation prevented gauge recording low tide. Can we now fill the gap in mean sea-level?

# Using citizen science to digitise 3 million hand-written tide-gauge data entries

**JOANNE WILLIAMS**, ANDREW MATTHEWS, ELIZABETH BRADSHAW & MORE THAN 3831 VOLUNTEERS

National Oceanography Centre, Liverpool, UK. joll@noc.ac.uk

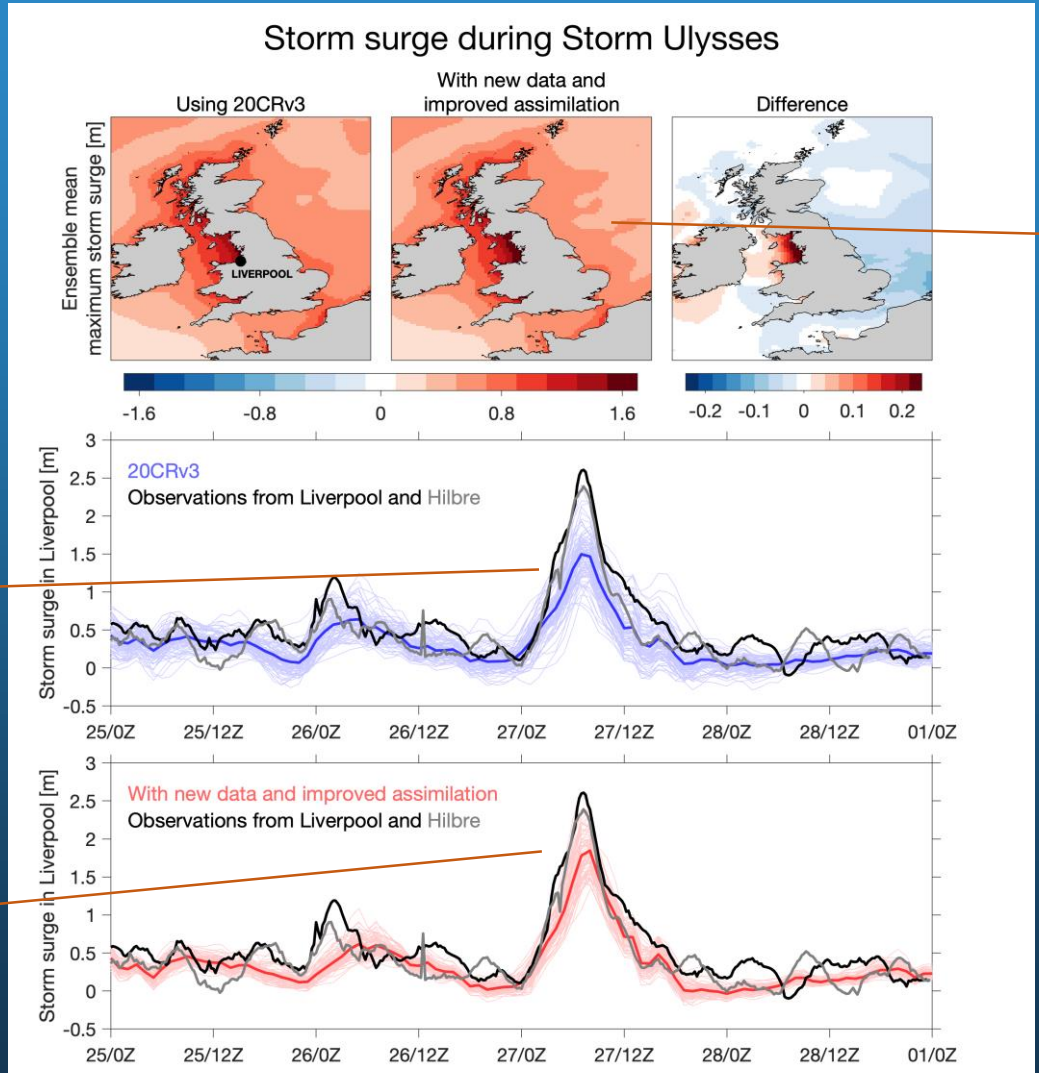


EGU General Assembly
23rd May 2021
EGU22-740

The new 15 minute data from 1857-1903 will allow us to almost double the storm surge extremes record at Liverpool, and for the first time validate it at nearby Hilbre.

Non-tidal residuals from Hilbre and Liverpool in good agreement with each other

Better agreement with the surge model based on new weather




*Thanks to Ed Hawkins, Kiki Kuijjer, Ivan Haigh et al*

New surge forecast based on updated weather reanalysis, using rescued pressure records from WeatherRescue.org citizen science project.

I'm a retired nurse researcher in Pennsylvania, US and I'm really enjoying entering this data. I'm curious about the men (I presume men or boys) that recorded this so long ago. Were they local children? … Thanks for letting us share in your data retrieval!

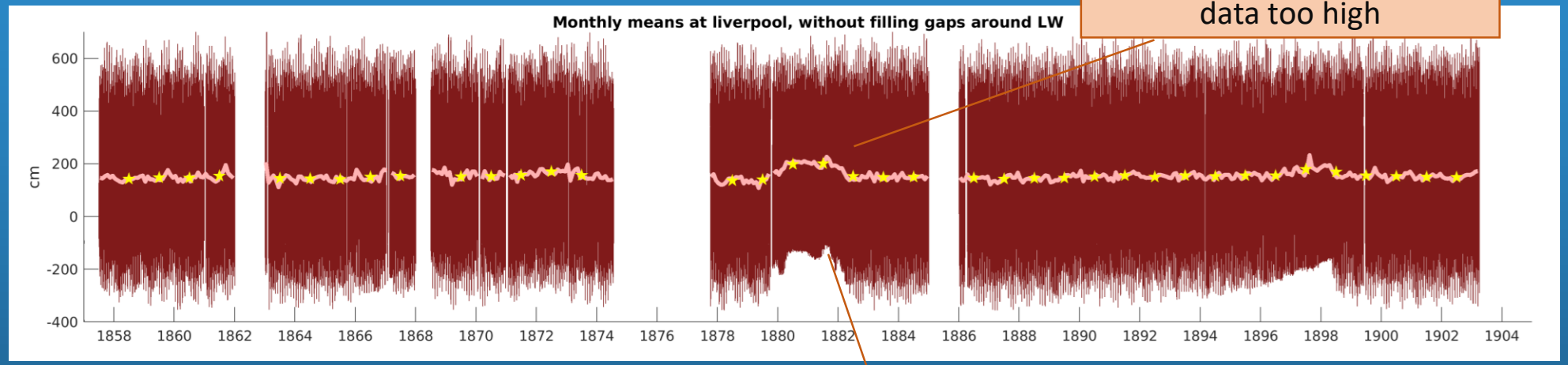
# Using citizen science to digitise 3 million hand-written tide-gauge data entries

**JOANNE WILLIAMS**, ANDREW MATTHEWS, ELIZABETH BRADSHAW & MORE THAN 3831 VOLUNTEERS

National Oceanography Centre, Liverpool, UK. joll@noc.ac.uk

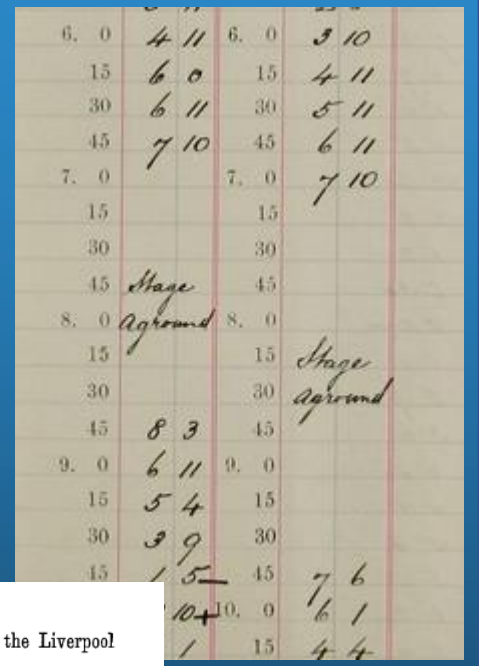


EGU General Assembly
23rd May 2021
EGU22-740

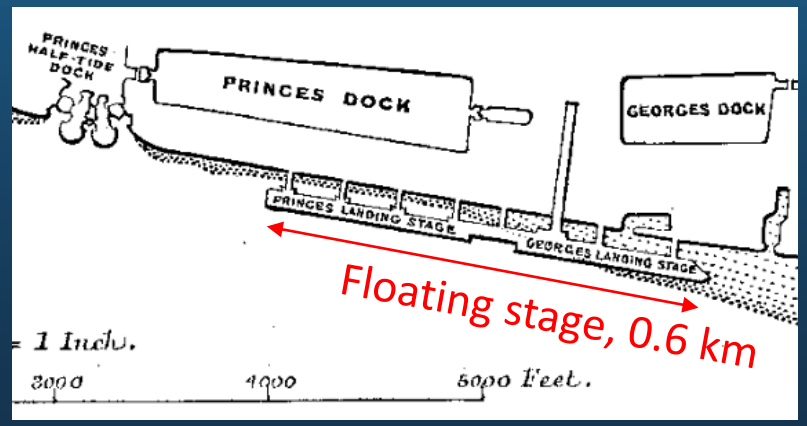



Monthly mean of recorded data too high

Siltation prevented gauge recording low tide.



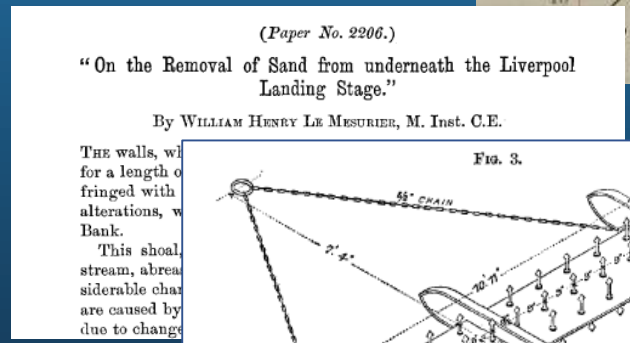

The gauge at Liverpool used the 600m landing stage which floated on the tide




Some years sediment built up under the stage, preventing it from reaching Low Water

Monthly means of available data would be incorrect, so there are gaps in PSMSL mean sea level record.

Now we have high frequency data, can we fill the gaps?

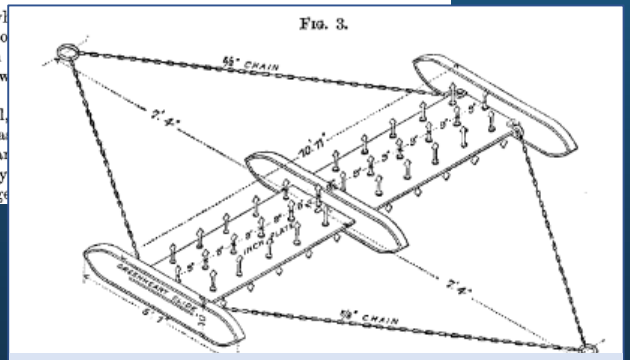


(*Paper No. 2206.*)

"On the Removal of Sand from underneath the Liverpool Landing Stage."

By WILLIAM HENRY LE MESURIER, M. Inst. C.E.

THE walls, wh
for a length o
fringed with
alterations, w
Bank.
This shoal,
stream, abrea
siderable char
are caused by
due to change




Harrow design to clear sediment

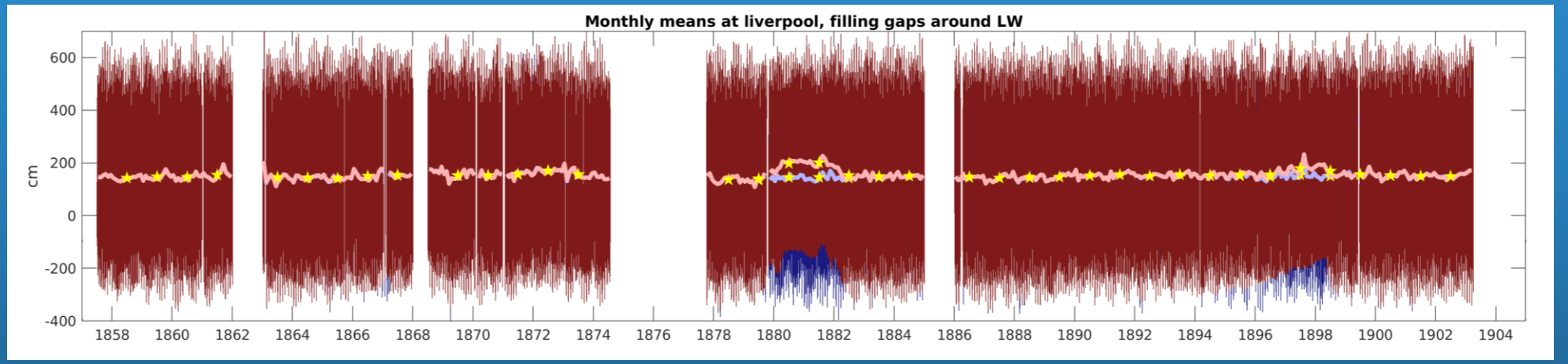
# Using citizen science to digitise 3 million hand-written tide-gauge data entries

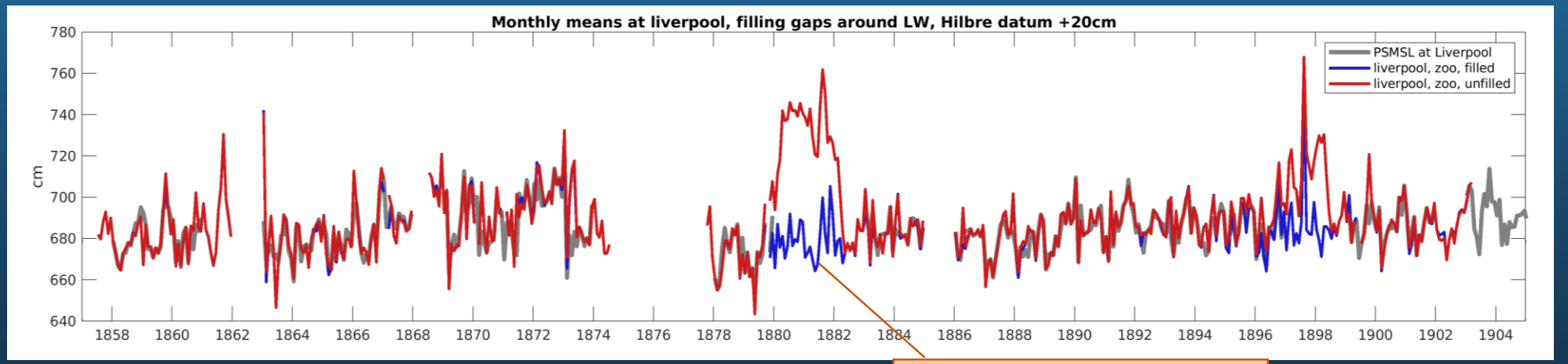
**JOANNE WILLIAMS**, ANDREW MATTHEWS, ELIZABETH BRADSHAW & MORE THAN 3831 VOLUNTEERS

National Oceanography Centre, Liverpool, UK. joll@noc.ac.uk



EGU General Assembly
23rd May 2021
EGU22-740




Monthly means after filling gaps at low water

To fill gaps at Low Water caused by landing stage grounding on sediment:

- Remove tidal prediction based on more complete data in other years.
- Interpolate non-tidal residual.
- Restore tidal prediction

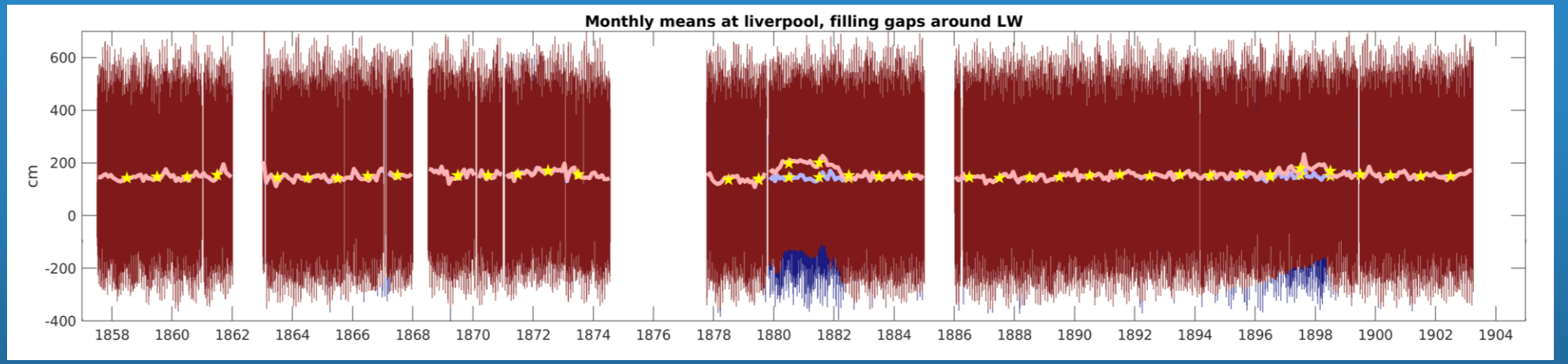
# Using citizen science to digitise 3 million hand-written tide-gauge data entries

**JOANNE WILLIAMS**, ANDREW MATTHEWS, ELIZABETH BRADSHAW & MORE THAN 3831 VOLUNTEERS

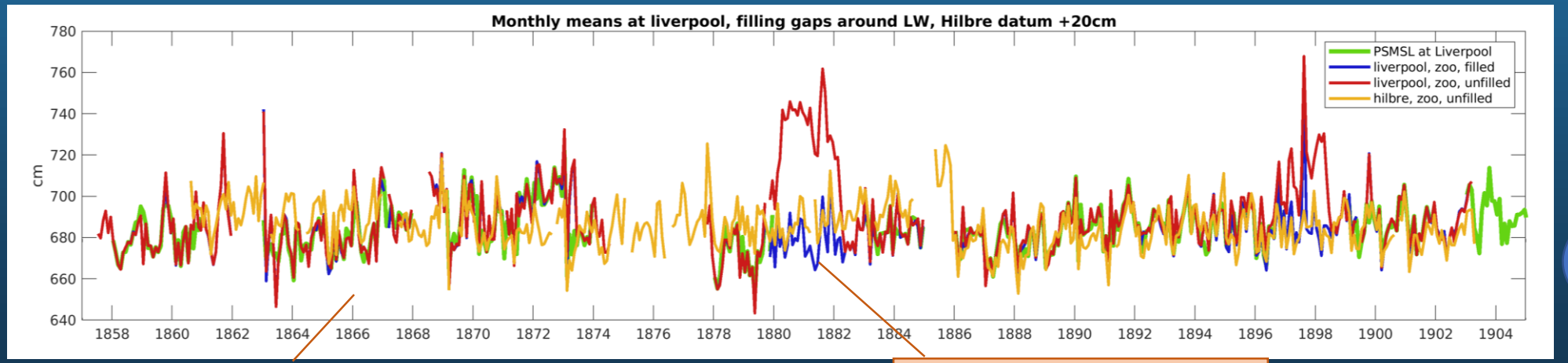
National Oceanography Centre, Liverpool, UK. joll@noc.ac.uk



EGU General Assembly
23rd May 2021
EGU22-740




Gap filling looks promising – further cross validation with Hilbre will be possible when quality control is complete.




QC not complete on earliest records – treat with caution!

New Liverpool monthly means after filling gaps are similar to Hilbre

Will there be another set after that? I confess I am a bit of a Zooniverse junkie

13:37–13:44 EGU22-740 On-site presentation

# Using citizen science to digitise 3 million hand-written tide-gauge data entries

Joanne Williams, Andrew Matthews, and Elizabeth Bradshaw

*How can you get sea-level data faster than one day at a time? Get it from the past!*

*The port of Liverpool is one of the world's longest sea-level records, but for the 1800s the only digital record is hand-calculated monthly mean data, which have many gaps. Hand-written ledgers contain high frequency (15 minute) records from 1853 to 1903, both at Liverpool and neighbouring Hilbre Island. In 2021, we coordinated over 3600 volunteers through the Zooniverse website to transcribe this data. At the time of writing this abstract, the transcription is nearing completion. From the newly digitised data we can examine whether tides in the Mersey have changed and reassess the frequency of rare storm surge events. We now understand the reason for the gaps in the Liverpool monthly mean sea-level, which are due to a dock fire and an intermittent siltation problem at low water, and may be able to use the Hilbre data to help fill them.*

*We report on the feasibility of this process for other transcription projects, the unusual quality control requirements for volunteer transcription, and present the newly restored data with 19th Century tides, storm surges and sea-level.*

How to cite: Williams, J., Matthews, A., and Bradshaw, E.: Using citizen science to digitise 3 million hand-written tide-gauge data entries, EGU General Assembly 2022, Vienna, Austria, 23–27 May 2022, EGU22-740, https://doi.org/10.5194/egusphere-egu22-740, 2022.