Big data: Challenges and solutions of archiving over 130,000 years of sea-level change

Patrick T. Boyden[1] and Sophie Williams[2]

Dublin, Ireland, 21-23 July 2019

The PALSEA (PALeo constraints on SEA level rise; pastglobalchanges.org/palsea) working group convened recently at Trinity College, Dublin. This meeting (pastglobalchanges.org/calendar/2019/127-pages/1821) was the first in the new phase (2019-2021) of the PAGES-INQUA project focusing on using ecological and chronological data to refine proxy-based reconstructions of sea level in light of recent developments in the discipline (e.g. Barlow et al. 2018; Shennan et al. 2018; Capron et al. 2019). The end goal of this work is to produce standardized sea-level databases. The meeting particularly focused on two large databases: HOLSEA (focusing from the Last Glacial Maximum (LGM) to present) and WARMCOASTS (focusing on the Last Interglacial; Fig. 1).

The HOLSEA project (Geographic variability of HOLocene SEA level; holsea.org) is concerned with determining the geographic variability of sea-level change in the last ~20,000 years. The project, led by Nicole Khan at the University of Hong Kong and funded by INQUA, seeks to better quantify how global sea level has changed since the LGM, and identify the driving trends of spatial variability in relative sea level over this time period. While the HOLSEA project has several aims, the specific goals of the breakout group were to discuss the data archiving process and to address any comments on the sea-level data template used within the HOLSEA database introduced by Khan et al. (2019).

User interaction with the database was discussed, with emphasis on challenges faced when entering poorly resolved sites. Well-resolved sites primarily use sedimentary indicators and are easy to enter. Poorly resolved sites, however, are located in the high latitudes, where geomorphic indicators are used and environmental variables are difficult to obtain. Other issues included accessibility to data, conflicting data interpretation, and problems with old data using outdated radiocarbon calibration curves.

The second half of the discussion focused on using the database to address challenges across interdisciplinary boundaries. Finally, the group discussed future uses of the database in addressing spatial variability of sea level, fingerprinting sources of ice melt, reconstructing paleo-shoreline changes, and integrating the database into glacial isostatic adjustment (GIA) models.

WARMCOASTS (Sea level and extreme waves in the Last Interglacial) is led by Alessio Rovere and funded by an ERC starting grant. The project focuses on the reconstruction of sea level and analysis of extreme waves during the Last Interglacial (LIG), approximately 125 kyr BP. Drawing on work by the HOLSEA group, the overarching goal of WARMCOASTS consists of advancing our current understanding of LIG sea-level changes and reducing uncertainties related to field evidence. The first phase of the project has begun with the development of

a global, open-access database: the World Atlas of Last Interglacial Shorelines (WALIS). This provides a standardized template for scientists from around the world to make their published datasets available for other scientists to utilize. More information can be found on the WARMCOASTS website (warmcoasts.eu).

During the WARMCOASTS breakout session, the database development team introduced the data management user interface. Participants discussed questions regarding the logistics of programing and managing the database. The current development team is small in number but enthusiastic about wider collaboration. A central theme of the session was a debate regarding whether to include only raw data or also to provide a non-quantifiable description of the corresponding data point. It was agreed that while raw data itself should be the main focus of the database, a short description of geological context would prove to be invaluable for those using the data in the future.

This was the first of three foreseen PALSEA meetings during this phase, and was successful in leading to more cohesion within the sea-level science community, particularly as we utilize more "big data" approaches in the discipline. The next workshop in 2020 is planned to be held at Columbia University and will focus on the recent advancements in GIA and ice-sheet models and their contribution to our understanding of paleo shorelines.

AFFILIATIONS
[1]MARUM - Center for Marine Environmental Science, University of Bremen, Germany
[2]Department of Environment and Geography, University of York, United Kingdom

CONTACT
Patrick T. Boyden: pboyden@marum.de

REFERENCES
Barlow NLM et al. (2018) Nat Geosci 11: 627-634
Capron E et al. (2019) Quat Sci Rev 219: 308-311
Hibbert FD et al. (2016) Quat Sci Rev 145: 1-56
Khan NS et al. (2019) Quat Sci Rev 220: 359-371
Kopp RE et al. (2009) Nature 462: 863-867
Pedoja K et al. (2011) Earth Sci Rev 108: 1-15
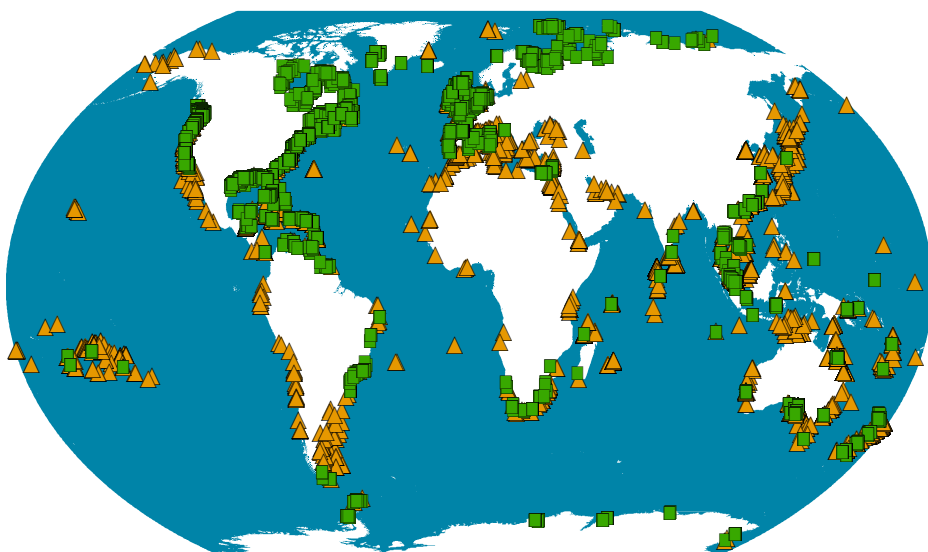Shennan I et al. (2018) Quat Sci Rev 201: 380-395

Figure 1: Overview of current paleo sea-level data within the HOLSEA (green squares) and WARMCOAST (orange triangles) databases. Initial work on the WARMCOAST database has been to combine data from previously published datasets as a foundation (Kopp et al. 2009; Pedoja et al. 2011; Hibbert et al. 2016).