



## Palaeo-sea-level and palaeo-ice-sheet databases: problems, strategies, and perspectives

André Düsterhus<sup>1,2</sup>, Alessio Rovere<sup>3,4</sup>, Anders E. Carlson<sup>5</sup>, Benjamin P. Horton<sup>6,7</sup>, Volker Klemann<sup>8</sup>, Lev Tarasov<sup>9</sup>, Natasha L. M. Barlow<sup>10</sup>, Tom Bradwell<sup>11</sup>, Jorie Clark<sup>5</sup>, Andrea Dutton<sup>12</sup>, W. Roland Gehrels<sup>13</sup>, Fiona D. Hibbert<sup>14</sup>, Marc P. Hijma<sup>15</sup>, Nicole Khan<sup>6</sup>, Robert E. Kopp<sup>16</sup>, Dorit Sivan<sup>17</sup>, and Torbjörn E. Törnqvist<sup>18</sup>

<sup>1</sup>National Oceanography Centre, Liverpool, L3 5DA, UK

<sup>2</sup>Institute of Oceanography, Center for Earth System Research and Sustainability (CEN), University of Hamburg, Bundesstraße 53, 20146 Hamburg, Germany

<sup>3</sup>MARUM, University of Bremen, & ZMT, Leibniz Center for Tropical Marine Ecology, Leobener Str., Bremen, Germany

<sup>4</sup>Lamont-Doherty Earth Observatory, Columbia University, P.O. Box 1000, 61 Route 9W, Palisades, New York, USA

<sup>5</sup>College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR 97331 USA

<sup>6</sup>Sea Level Research, Department of Marine and Coastal Science and Institute of Earth, Ocean and Atmospheric Sciences, Rutgers University, New Brunswick, NJ 08901, USA

<sup>7</sup>Earth Observatory of Singapore and Asian School of the Environment, Nanyang Technological University, 639798, Nanyang, Singapore

<sup>8</sup>Dep. Geodesy and Remote Sensing, German Research Centre for Geosciences GFZ, Potsdam, Germany

<sup>9</sup>Memorial University of Newfoundland, St. John's, NL, Canada

<sup>10</sup>Department of Geography, Durham University, South Road, Durham, DH1 3LE, UK

<sup>11</sup>Biological & Environmental Sciences, University of Stirling, Stirling, FK9 4LA, Scotland, UK

<sup>12</sup>Department of Geological Sciences, University of Florida, Gainesville, Florida, USA

<sup>13</sup>Environment Department, University of York, Heslington, York, YO10 5NG, UK

<sup>14</sup>Ocean and Earth Science, National Oceanography Centre, University of Southampton, Southampton, SO14 3ZH, UK

<sup>15</sup>Department of Applied Geology and Geophysics, Deltares, Utrecht, the Netherlands

<sup>16</sup>Department of Earth & Planetary Sciences, Rutgers Energy Institute, and Institute of Earth, Ocean, and Atmospheric Sciences, Rutgers University, New Brunswick, NJ, USA

<sup>17</sup>Department of Maritime Civilizations, Leon Charney School of Marine Sciences and Recanati Institute of Maritime Studies (RIMS), University of Haifa, Haifa, Israel

<sup>18</sup>Department of Earth and Environmental Sciences, Tulane University, 6823 St. Charles Avenue, New Orleans, Louisiana 70118-5698, USA

*Correspondence to:* André Düsterhus (andre.duesterhus@uni-hamburg.de)

Received: 22 May 2015 – Published in Clim. Past Discuss.: 25 June 2015

Revised: 16 February 2016 – Accepted: 7 March 2016 – Published: 11 April 2016

**Abstract.** Sea-level and ice-sheet databases have driven numerous advances in understanding the Earth system. We describe the challenges and offer best strategies that can be adopted to build self-consistent and standardised databases of geological and geochemical information used to archive palaeo-sea-levels and palaeo-ice-sheets. There are three phases in the development of a database: (i) measurement, (ii) interpretation, and (iii) database creation. Measure-

ment should include the objective description of the position and age of a sample, description of associated geological features, and quantification of uncertainties. Interpretation of the sample may have a subjective component, but it should always include uncertainties and alternative or contrasting interpretations, with any exclusion of existing interpretations requiring a full justification. During the creation of a database, an approach based on accessibility, transparency,

trust, availability, continuity, completeness, and communication of content (ATTAC<sup>3</sup>) must be adopted. It is essential to consider the community that creates and benefits from a database. We conclude that funding agencies should not only consider the creation of original data in specific research-question-oriented projects, but also include the possibility of using part of the funding for IT-related and database creation tasks, which are essential to guarantee accessibility and maintenance of the collected data.

## 1 Introduction

The rapid acquisition of palaeoclimate data and the development of strategies to assimilate these data into models has resulted in a growing need for open-access and user-friendly databases with the goal of machine readability (Overpeck et al., 2011). Within the palaeo-sea-level and palaeo-ice-sheet communities, there is the further requirement of standardisation (Hijma et al., 2015). These communities use field data to reconstruct the elevation of past sea levels and the dimensions and extent of former ice sheets. As an example of assimilation of data into models, databases of sea-level index points have constrained model estimates of the rates of glacial isostatic adjustment (GIA) during and following the last deglaciation (e.g. Milne et al., 2005; Bradley et al., 2011; Engelhart et al., 2011; Whitehouse et al., 2012; Peltier et al., 2015; Roy and Peltier, 2015). The results from these studies have contributed, in turn, to estimating current rates of ice-sheet mass loss and sea-level rise from geodetic observations (Vaughan et al., 2013). Other databases have been used to assess the magnitude of the sea-level highstand during the last interglacial period (Kopp et al., 2009; Dutton and Lambeck, 2012) and improve our understanding of global ocean volume and earth dynamic topography during the Pliocene (Rowley et al., 2013; Rovere et al., 2014, 2015). Likewise, the worldwide timing of the Last Glacial Maximum (e.g. Clark et al., 2009) and global deglaciation of valley glaciers (e.g. Shakun et al., 2015) has been determined from ice-sheet databases.

The generation of databases of past sea-level changes began with Daly (1934) and Godwin (1940), with early examples of reconstructing temporal changes in former ice-sheet margins by Prest et al. (1968) and Bryson et al. (1969). The need for standardisation among studies as new sea-level data emerged was recognised and implemented by International Geoscience Programme (IGCP) projects, starting with IGCP Project 61 in 1974 (van de Plassche, 1986). Subsequent IGCP projects produced Holocene databases in the United Kingdom (Shennan and Horton, 2002), the US Atlantic coast (Engelhart and Horton, 2012), South America (Milne et al., 2005), and elsewhere (Khan et al., 2015). Several recent studies have constructed deglacial databases of ice-sheet retreat, but they have used different criteria and approaches to data assimilation (e.g. Dyke, 2004; Clark et al., 2009; Tarasov

et al., 2012; Briggs et al., 2014; Hughes et al., 2016; Stokes et al., 2015; Stroeve et al., 2015).

The process of setting up a sea-level or ice-sheet database can be divided into three phases: (i) measurement, (ii) interpretation, and (iii) database creation. In this paper, we build on the results of PALSEA (PALeo constraints on SEA level rise; Siddall et al., 2010) workshops over the last 8 years to report the main challenges identified for each phase and the possible solutions that can be adopted.

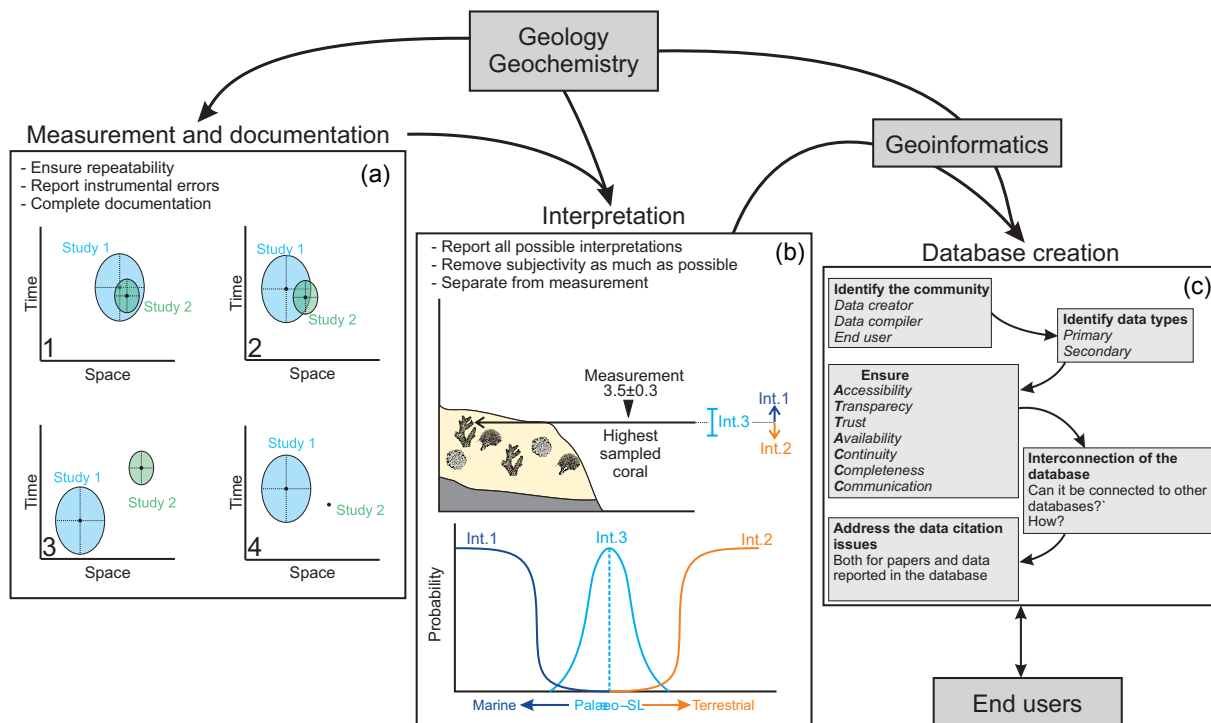
## 2 Measurements

A common denominator of palaeo-sea-level and palaeo-ice-sheet data is that they originate from two types of direct measurements. Field measurements are taken to determine the position, location, and elevation of a particular feature (e.g. a fossil coral or a glacial deposit). Meta data, such as cross sections and photographs, may also be used to illustrate the local geological and geomorphological context. Laboratory measurements include establishing the age of a feature (e.g. a <sup>14</sup>C or cosmogenic surface exposure age) which was sampled in the field. Sample information on location, elevation, and shielding for cosmogenic surface exposure ages is critical for recalculation of ages as inferred production rates change (Balco, 2011).

Any measurement of palaeo-sea-level and palaeo-ice-sheet data needs clearly specified measures of uncertainties. The scientific value of the data is maximised if uncertainties are reduced, but missing information often exacerbates difficulties in quantifying uncertainties. For example, uncertainties related to the elevation of a sea-level index point are potentially large if the original study did not indicate the tidal or geodetic data to which the elevation is referenced (van de Plassche, 1986). Elevation errors, which greatly affect palaeo-relative-sea-level (RSL) calculations, can be avoided by employing state-of-the-art GPS and levelling techniques (e.g. Muhs et al., 2011; Rovere et al., 2015). Despite this, high-accuracy GPS systems are to date seldom applied to measure Quaternary and Pliocene sea-level proxies. Although the laboratory error is often indicated as part of laboratory procedures, this is not always the case with instrumental errors in a field measurement.

Ideally, multiple studies measuring and interpreting the same proxy should have overlapping uncertainty ellipses (cases 1 and 2 in Fig. 1a). Unfortunately, there are many examples where measurements do not overlap (3 in Fig. 1a) or cannot be realistically compared due to the lack of details on measurement techniques or details on interpretation (4 in Fig. 1a). In the worst case, some studies may fail to report the error and cannot be compared. Incomplete data limit the longevity of some data, requiring new studies to remeasure the same proxies.

Measurement of palaeo-sea-level and palaeo-ice-sheet data can either be obtained by direct field or laboratory activ-



**Figure 1.** (a) The effect of different measurement and documentation of time and space of the same sea-level indicator in two different studies. (b) Different interpretations for the sea-level indicator: a deposit containing fossil corals. (c) ATTAC<sup>3</sup> approach to database creation. See text for details.

ities, or derived from a previous publication and inserted in the database. In all cases, the transfer of information should be objective and complete, reporting only what can be read in the original publication and/or what is measured in the field, with no further interpretation. An important goal for the future is for different communities to agree on standardised measurements and data reporting norms (e.g. Hijma et al., 2015). Precision of terminology is vital to avoid misinterpretations of field and laboratory measurements (Shennan et al., 2015). This will facilitate seamless interfacing with database systems for archiving and further analysis. Palaeo-sea-level and palaeo-ice-sheet databases need to include standardised documentation of fundamental data fields:

**Position** (i.e. geographical location and elevation or depth), referred to a specific sea-level datum and, if available, the positioning techniques applied;

**Age** including laboratory identification number, details on the dating technique used, and ideally the raw data;

**Description of the feature** including metadata and images to complement the quantitative information;

**Quantification** of measurement uncertainties.

### 3 Interpretation

Once measured, field and laboratory data are interpreted to reconstruct the palaeo-sea-level and the spatial and temporal extent of the palaeo-ice-sheet. Commonly it is the interpretation that will be most interesting for the final users, who may not be experts but need to compare the reconstructions with independent estimates, such as model predictions.

There is often a subjective component to the interpretation of field data. In Fig. 1b, we show fossil corals, a typical example of a sea-level indicator. An objective assessment of the coral age can be determined using U-series techniques (e.g. applying the template used in Dutton and Lambeck, 2012). The position of the deposit relative to a tidal or geodetic datum can be measured with appropriate accuracy. The taxonomy of the sample can be reported, which should include information on the benthic assemblage and its relation to sea level and geological and sedimentological properties. The subjectivity relates to the interpretation of the palaeowater depth (i.e. relation to sea level) of the coral. One possible interpretation following investigation of the depth distribution of corals in the deposit is that the corals are in situ (e.g. in living position) and sea level at the time of deposition was somewhere above the measured elevation (i.e. it is a lower limiting data point, Int. 1 in Fig. 1b). Another interpretation could be that the corals are allochthonous, and instead represent a storm deposit. In this case, it is only possible to in-

fer that the deposit represents the top of a marine sequence and that the palaeo-sea-level was located below the measured elevation of the deposit (i.e. its upper limiting data point, Int. 2 in Fig. 1b). A final interpretation may instead recognise elements (e.g. microatolls, intertidal geological facies within the deposit) that tie the deposit to the palaeo-sea-level around the measured elevation within an uncertainty (i.e. identifying reference water level and indicative range (Shennan, 1986; van de Plassche, 1986; Horton et al., 2000), Int. 3 in Fig. 1b). Whenever controversial interpretations such as those summarised above exist, a database should document all of them. If one interpretation is more likely than the others, or is supported by independent studies, this information should be inserted in the database within the metadata. Issues may emerge in the interpretation of laboratory data, such as the use of different calibration curves to establish the age of an indicator. The interpretation of data can be subject to changes with scientific advances. As an example, old  $^{14}\text{C}$  ages or cosmogenic surface exposure ages can be recalibrated following the availability of new calibration curves or calibration schemes or new production rates and scaling models, respectively. But the possibility of recalibrating these measurements depends on the presence of primary data, such as  $\delta^{13}\text{C}$  measurements, description of the dated material, sample thickness, etc. (Balco, 2011; Törnqvist et al., 2015). In principle, if measurement data are present in a database, obtaining secondary data from new interpretations can be streamlined relatively easily.

Uncertainties of sea-level and ice-sheet indicators are usually treated as Gaussian distributions, with the exception of limiting data that only provide information on maximum or minimum sea level (Int. 1 and Int. 2 in Fig. 1b). In the case of Gaussian uncertainties, the uncertainty of the interpretation can be combined with the uncertainty of the measurement (dashed line in Fig. 1b) using the root mean square error formula assuming the uncertainties are independent; more complicated uncertainties may require Monte Carlo sampling. As understanding of habitat distribution for marine species or coastal facies increases, and more consideration is given to the physical processes that perturb sample elevation over time, an increasing amount of data will use more accurate uncertainty distributions that extend beyond the Gaussian approximation. We recommend recording multiple percentiles of these non-Gaussian distributions to reflect not only the width of the uncertainty but also the shape of its probability distribution.

Palaeo-sea-level and palaeo-ice-sheet databases that incorporate interpretations must therefore be

**Flexible** to take into account the fact that, although the measurement must be unequivocal, interpretation of the data can be multiple and vary or evolve over time;

**Consistent** in the reporting of interpretations and uncertainties of data.

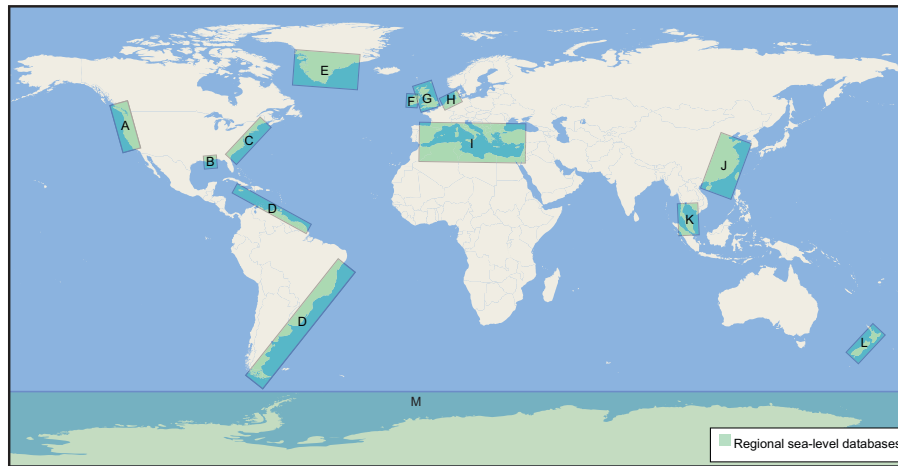
## 4 Database creation

A database is primarily a collection of data records and secondarily a platform for exchange of data and information. The process of creating a database must necessarily start from the identification of the agents that will interact with it. Data creators provide the original data sets and should carry out their work with databases in mind. In palaeoclimate sciences, these are usually geologists and geochemists, who carry out the main part of the measurement and interpretation process. Data compilers collect data from different sources and, if necessary, reinterpret it. Measurement and interpretation constitute the backbone of every palaeo-sea-level and palaeo-ice-sheet database, but there are other key elements to be considered, which we summarise under the ATTAC<sup>3</sup> acronym (Fig. 1c): accessibility, transparency, trust, availability, continuity, completeness, and communication of content.

**Accessibility** is a challenge due to the heterogeneity of the user communities. A majority of published databases today use a spreadsheet format, which is easy to access for most users. However, some information (e.g. images or non-Gaussian uncertainty estimates) is more simply presented in relational databases. Furthermore, relational databases enable different presentation formats for different end-user communities.

**Transparency** is critical in interdisciplinary research fields. Scientists must trust each other on the applied methodology, but at the same time they have to be able to understand the applied procedures. As a database creator and compiler cannot know all future users of the data and the fields in which they are applied, the database description must be as detailed as possible. The description should include appropriate metadata and use standardised language and comments in data fields. Indicating the quality of each data field in understandable formats will help the end user to make appropriate use of the data (Düsterhus and Hense, 2014).

**Trust** is built by database compilers sharing credit with the scientists delivering the data (Costello, 2009; Kattage et al., 2014). Data creators and compilers are confronted with the risk that their original publications are no longer cited when their data are included in a larger citable database and thus will not gain credit under current performance metrics, such as the H-index (Hirsch, 2005). To ensure the availability of high-quality data sets in the future, data creators need to be given appropriate credit. Trust of a database requires consistent data quality and transparently applied procedures, and a consistent and trustworthy host. It also requires effective software design for the database and within the data processing. Ideally, the code should be openly available and well documented.



**Figure 2.** Published regional sea-level databases that follow (where appropriate) IGCP guidelines: (A) Pacific (Engelhart et al., 2015; Reynolds and Simms, 2015), (B) Gulf (Hijma et al., 2015), and (C) Atlantic (Engelhart and Horton, 2012) coasts of North America; (D) Caribbean and South America (Milne et al., 2005); (E) Greenland (Long et al., 2011); (F) Ireland (Brooks and Edwards, 2006); (G) the UK (Shennan and Horton, 2002); (H) northwest Europe (Vink et al., 2007); (I) the Mediterranean (Vacchi et al., 2014, 2016); (J) China (Zong, 2004), (K) Malay Peninsula (Horton et al., 2005); (L) New Zealand (Clement et al., 2016); and (M) Antarctica (Briggs and Tarasov, 2013).

**Availability** of a database for the long term requires long-term funding (see below). Today, most databases are attached to journal articles as a spreadsheet in the supplement. This ensures persistence, but no database maintenance and/or upgrade is possible for most journals.

**Continuity** of updating is important to stay relevant and reflect the changing interpretations of the data. To allow cite-ability of the database (e.g. with digital object identifiers (DOIs); Paskin, 2005; Quadt et al., 2012), version control is essential. Furthermore, the use of unique and persistent identifiers, such as the International Geo Sample Number (IGSN) that is currently used for geological samples, should be encouraged to ensure that over different update cycles a data point can be uniquely referenced by scientists.

**Completeness** of the database is important, especially in the context of uncertainties (Hijma et al., 2015). Even when the basic elements (like position, age, and elevation) are complete, for many applications they are of limited use when associated uncertainties are not clearly indicated or defined.

**Communication** of the content, for example through interfaces for visualisation software, helps to open the database for new audiences. Advanced visualisation approaches (e.g. Rovere et al., 2012; Unger et al., 2012) require standardised protocols for data extraction and consistent data types. These properties have to be determined in the design phase of the database; thus it is important to consider its applications right from the beginning.

## 5 The community structure

Any database should be aimed at serving a community of end users, who extract content for further analysis, and give feedback on specific needs regarding data sets or analyses. Databases should be centralised and interconnected via the Internet in order to reach the maximum possible number of end users, with the widest possible geographic distribution. The data are more likely to be used if the end users have a unique access point for the data sets, such as a WebGIS portal. In the geological domain, there are large initiatives to build data repositories, which are already well established and used by scientists worldwide. Two examples are the NOAA World Data Center for Paleoclimatology (Wahl et al., 2010) and PANGAEA (<http://www.pangaea.de/>). Some journals link PANGAEA databases to online versions of associated papers.

Most funding agencies require that data collected in the framework of a project be archived and made available through data repositories. This is achieved through a “data management plan” (National Science Foundation of the United States) or the “open data policy” (European Union), which requests that the project leaders state where they plan to store the data collected within their project. Currently, a researcher working on sea-level and/or ice-sheet databases only has the choice to store the new data sets in different repositories, which might have the effect of dispersing the data across several repositories, decentralising data storage (see example in Fig. 3).

In the framework of a single research project, the data creator is also a data compiler, and often the first end user. It is, therefore, necessary to ensure that the data sets collected in





## Data Management Plan

Primary Investigator: Researcher Name

Institution: University

Project: Sea level data management plan example

Co-PIs: Researcher names

NSF Division: OCE Solicitation Info: Marine geology and geophysics Submission Date: 12/24/2015

Overview: We will produce new data in the format of field notes, photos of field sites, GPS data, GIS datasets and databases. At each site we will collect samples that will be dated with U-series techniques.

Data description: We expect to produce information on paleo sea levels measured in the field and we will revise information as necessary, on sites published in literature.

Description of existing data and samples: We will use existing published data on sea levels in the area of interest. We will collect all the information in a geodatabase built with GIS software, from which, at the end of the project, we will extract the information to be submitted to the repositories listed hereafter.

Data analysis summary: Samples collected in the field will be dated using U-series.

Includes field work? Yes

Description of field work: We will collect elevation and stratigraphic data in the field with a differential GPS receiver. Cameras of the researchers will be synchronized with their handheld GPS time to geotag field photos. Field notes will be digitized at the end of each day of field survey and stored in PDF format.

### **Expected data product #1**

Data type: Observational

Responsible investigator: Researcher

Product description: Geologic samples from units surveyed in the field. Sample metadata will include pictures, Lat/Lon coordinates, description of the facies and sketch of the sampling location complete with GPS elevations.

Intended repository: SESAR

Timeline for data release: Two Years from acquisition/analysis

### **Expected data product #2**

Data type: Observational

Responsible investigator: Researcher

Product description: GPS coordinates and GPS raw data of sampling localities, measured with DGPS. Pictures, text description, field sketches and associated ISGN numbers will be used as metadata.

Intended repository: UNAVCO

Timeline for data release: Two Years from acquisition/analysis

### **Expected data product #3**

Data type: Analytical

Responsible investigator: Researcher

Product description: Results of U-series analyses. Metadata will include photos of samples, description of facies, lat/lon coordinates and ISGN numbers associated.

Intended repository: EarthChem

Timeline for data release: Two Years from acquisition/analysis

### **Expected data product #4**

Data type: Observational

Responsible investigator: Researcher

Product description: New site stratigraphies as well as stratigraphies re-evaluated or re-measured from literature data as necessary.

Intended repository: GeoStrat

Timeline for data release: Two Years from acquisition/analysis

**Figure 3.** Example of a data management plan (DMP) for a project on Pleistocene sea-level markers obtained with the IEDA (Interdisciplinary Earth Data Alliance, <http://www.iedadata.org/>) DMP toolbox. Note that, to correctly store sea-level data, at least four independent repositories are needed.

**Table 1.** List of published global sea-level databases that follow (where appropriate) the IGCP format. Formats in which the data are provided: spreadsheets (R) and interactive interfaces that allow the visualisation, extraction, or download of data (I).

Description	Accessibility
Dutton and Lambeck (2012) Compilation of last interglacial coral U–Th age data, elevation data, and associated sample information. The first worksheet of the Excel file contains the data and calculated ages and elevations that have been normalised to common decay constants and elevation benchmarks, respectively; the second worksheet contains definitions of column headings and data units; the third worksheet contains a lookup table for data sources listed by number in the first worksheet. Some entries in the database are annotated by comment fields to denote supplemental information for data or calculations not included in the original publications.	Annexed to publ. (R)
Klemann et al. (2013) Storage of different accessible compilations in relational database system PostgreSQL. Contains the regional databases A, B, C, and D shown in Fig. 2 and further data mainly from published compilations or grey literature. Access via visualisation and analysis software SLIVISU (beta version) or direct access (password protected).	Online (I), on request
Khan et al. (2015) Compilation of global Holocene relative sea-level data. Each database entry includes location, sea level, sea-level error, age, and age error, as well as the original source of publication.	Annexed to publ. (R)
Kopp et al. (2009) Multi-proxy database of last interglacial index and limiting relative sea-level index points. A legend worksheet defines column headings and data units.	Annexed to publ. (R)
Kopp et al. (2016, <a href="https://www.ncdc.noaa.gov/paleo/study/19823">https://www.ncdc.noaa.gov/paleo/study/19823</a> ) Database of Common Era (last 3000 years) relative sea-level data. Each database entry includes location, sea level, sea-level error, age, and age error, as well as the original source publication. There is a front page of definitions of column headings and data units.	Online (R)
Pedaja et al. (2014) Spreadsheets containing information on shoreline analysis from Holocene to Miocene highstands. Regarding ages, only stratigraphic units are given.	Annexed to publ. (R)
Rovere et al. (2012, <a href="http://pliomax.org/pliowiki/index.php/RSLmap">http://pliomax.org/pliowiki/index.php/RSLmap</a> ) RSLmap is a visualisation tool for RSL markers, which allows the display and querying of a database of published or user-submitted relative sea-level data points.	Online (I)

the framework of a single project have a standardised structure and are available to other end users.

A significant concern regarding the maintenance of a healthy research community is appropriate crediting of authorship. How does an end user using thousands of data points from dozens of source publications provide appropriate credit? Journals often allow for only a limited number of citations, and often the citation credit goes to the data compiler, who created the review database, and not to the data creator. If the question above is not addressed, the long-term result will be that data creators will have no incentive to support the inclusion of their work in a centralised database. This issue must be addressed by journal editors. In some cases, editors have made exceptions to standard journal length rules in order to include all the original papers in the reference list (e.g. Khan et al., 2015). Alternatively, some journals allow longer, online-only papers with space for a full reference list (e.g. Kopp et al., 2016). A number of sea-level

databases have been produced in the framework of single research projects (Table 1). In general, there are two formats in which the data are provided: data repositories in the form of spreadsheets (R) and interactive interfaces that allow the visualisation, extraction, or download of data (I). In Fig. 2 we show the geographic coverage for a number of databases representing late-glacial and Holocene RSL data which were compiled from different original studies following, where appropriate, the IGCP guidelines. Each index point has a defined location, age, elevation relative to former sea level, and appropriate accounting of errors (details to the databases in Table 2).

**Table 2.** Description of regional databases presented in Fig. 2. For details see Table 1.

Region	Description	Accessibility
A	Engelhart et al. (2015); Reynolds and Simms (2015) ( <a href="http://sealevel.marine.rutgers.edu/">http://sealevel.marine.rutgers.edu/</a> ). Deglacial sea-level database for the Pacific coast of central North America.	Online (R)
B	Hijma et al. (2015, <a href="https://www.ncdc.noaa.gov/paleo/study/16361">https://www.ncdc.noaa.gov/paleo/study/16361</a> ). Pilot database intended as an initial release of Holocene geological relative sea-level data that have been compiled according to a recently developed protocol (Hijma et al., 2015). The database is provided in two versions: a complete version that consists of 77 variables and that includes all the underlying data, as well as a processed version with only the 11 most critical variables. It is anticipated that this latter version will be adequate for most users, while the former provides a full documentation for those who wish to carry out more detailed analyses.	Online (R)
C	Engelhart and Horton (2012, <a href="http://sealevel.marine.rutgers.edu/">http://sealevel.marine.rutgers.edu/</a> ). Holocene sea-level database for the Atlantic coast of the United States.	Online (R)
D	Milne et al. (2005). Deglacial sea-level compilation for the Caribbean and South American Atlantic coast that was compiled for a regional GIA study.	Appendix to publ.
E	Long et al. (2011). Compilation and own investigations mainly of Holocene isolation basins for southern Greenland.	Table in publ.
F	Brooks and Edwards (2006, <a href="http://www.naturalscience.tcd.ie/SL_Database.php">http://www.naturalscience.tcd.ie/SL_Database.php</a> ). Compilation of sea-level data of Ireland, which contains detailed spreadsheet and additional information on webpage.	Online (R)
G	Shennan and Horton (2002). The database covers Great Britain (England, Scotland, and Wales) and has around 2250 entries. It exists as an Access database; interfaces to convert the information to Excel spreadsheets for regional compilations are available.	From author
H	Vink et al. (2007). Compilation of available data of Belgium, the Netherlands, and Germany of the Channel and southern North Sea of about 380 SLIs which are listed in an appendix to the publication.	Appendix to publ.
I	Vacchi et al. (2014, 2016, <a href="http://www.medflood.org/results-2/webgis/">http://www.medflood.org/results-2/webgis/</a> ). Published Pleistocene and Holocene sea-level data in the Mediterranean Sea collected by the INQUA MEDFLOOD project.	Online (I)
J	Zong (2004). Compilation of SLIs covering SE coast of China. There are only a few attributes listed.	Appendix to publ.
K	Horton et al. (2005, <a href="http://sealevel.marine.rutgers.edu/">http://sealevel.marine.rutgers.edu/</a> ). Holocene sea-levels database of Malay–Thai Peninsula, Southeast Asia.	Online (R)
L	Clement et al. (2016). Compilation of Holocene sea-level data of New Zealand.	Table in publ.
M	Briggs and Tarasov (2013). Compilation of late-glacial and Holocene sea-level data of Antarctica.	Table in publ.

## 6 Concluding remarks

The discussions of the PALSEA community on sea-level and ice -heet databases can be framed around the following points:

- Any set-up of sea-level or ice-sheet databases must be divided into
  - measurement,
  - interpretation,
  - database creation.
- Storage of measurements should include position, age, description of geological features, and quantification of uncertainties. All must be described as objectively as possible with relevant metadata.



3. Interpretation of geological data will retain a subjective component, but it should always include uncertainties and include all the possible interpretations.
4. When creating a database, all the aspects related to the ATTAC<sup>3</sup> approach must be taken into account.
5. The community structure that creates and benefits from a database must be considered, and the needs and concerns of each part of the community must be respected.

There remains the need for a centralised database structure for the sea-level and ice-sheet communities. Despite this need, dedicated funding for “user-friendly”, field-specific database creation is rarely available because funding mostly prioritises projects that follow the classic hypothesis-driven research approach. Data management is often restricted to archiving at a general level. The tasks of database creation, maintenance, and guarantee of accessibility are limited to single projects, and the possibility to hire ad hoc personnel (e.g. experts in geoinformatics) to fulfil these requirements is often disregarded by funding agencies. We favour interdisciplinary research collaborations focusing on field-specific database development and maintenance, including projects that amalgamate and reanalyse published data sets into new databases. These new databases enhance the legacy of monetary investments originally made to collect sea-level and ice-sheet data. Many of the aspects discussed in this paper will also be valid for other types of geological data and may be of interest to additional geoscientific communities.

**Acknowledgements.** This paper has been written in the framework of PALSEA, the PALeo constraints on SEA level rise project, sponsored by PAGES and INQUA. This publication was planned at the PALSEA workshop in Lochinver, Scotland, organised by Antony Long and Natasha Barlow (Durham University). Publication costs were covered by Tulane’s Vokes Geology Fund (Torbjörn E. Törnqvist). The authors wish to thank the following funding sources: Natural Environmental Research Council under the numbers NE/I008365/1 (André Düsterhus) and NE/I008624/1 (W. Roland Gehrels); Institutional Strategy of the University of Bremen (German Excellence Initiative, Alessio Rovere); the US National Science Foundation (NSF) grant OCE1458904 (Benjamin P. Horton, Robert E. Kopp), grant OCE-1502588 (Torbjörn Törnqvist) and award no. 1443037 (Andrea Dutton); Global Change Program (EAR 1304909, Jorie Clark); and the German Climate Modeling Initiative (PALMOD, Volker Klemann).

Edited by: D. Fleitmann

## References

Balco, G.: Contributions and unrealized potential contributions of cosmogenic-nuclide exposure dating to glacier chronology, 1990–2010, *Quaternary Sci. Rev.*, 30, 3–27, doi:10.1016/j.quascirev.2010.11.003, 2011.

- Bradley, S. L., Milne, G. A., Shennan, I., and Edwards, R.: An improved glacial isostatic adjustment model for the British Isles, *J. Quat. Sci.*, 26, 541–552, doi:10.1002/jqs.1481, 2011.
- Briggs, R. D. and Tarasov, L.: How to evaluate model-derived deglaciation chronologies: a case study using Antarctica, *Quaternary Sci. Rev.*, 63, 109–127, doi:10.1016/j.quascirev.2012.11.021, 2013.
- Briggs, R. D., Pollard, D., and Tarasov, L.: A data-constrained large ensemble analysis of Antarctic evolution since the Eemian, *Quaternary Sci. Rev.*, 103, 91–115, doi:10.1016/j.quascirev.2014.09.003, 2014.
- Brooks, A. and Edwards, R.: The Development of a Sea-Level Database for Ireland, *Irish J. Earth Sci.*, 24, 13–27, doi:10.3318/ijes.2006.24.1.13, 2006.
- Bryson, R. A., Wendland, W. M., Ives, J. D., and Andrews, J. T.: Radiocarbon Isochrones on the Disintegration of the Laurentide Ice Sheet, *Arctic Alpine Res.*, 1, 1–14, doi:10.2307/1550356, 1969.
- Clark, P. U., Dyke, A., Shakun, J. D., Carlson, A. E., Clark, J., Wohlfarth, B., Hostetler, S., Mitrovica, J. X., and McCabe, A.: The Last Glacial Maximum, *Science*, 325, 710–714, doi:10.1126/science.1172873, 2009.
- Clement, A. J., Whitehouse, P. L., and Sloss, C. R.: An examination of spatial variability in the timing and magnitude of Holocene relative sea-level changes in the New Zealand archipelago, *Quaternary Sci. Rev.*, 131, 73–101, doi:10.1016/j.quascirev.2015.09.025, 2016.
- Costello, M. J.: Motivating Online Publication of Data, *BioScience*, 59, 418–427, doi:10.1525/bio.2009.59.5.9, 2009.
- Daly, R. A.: The changing world of the ice age, Yale University Press, New Haven, doi:10.1017/s001675680009422x, 1934.
- Düsterhus, A. and Hense, A.: Automated quality evaluation for a more effective data peer review, *Data Sci. J.*, 13, 67–78, doi:10.2481/dsj.14-009, 2014.
- Dutton, A. and Lambeck, K.: Ice Volume and Sea Level During the Last Interglacial, *Science*, 337, 216–219, doi:10.1126/science.1205749, 2012.
- Dyke, A. S.: An outline of North American Deglaciation with emphasis on central and northern Canada, in: *Quaternary Glaciations – Extent and Chronology: Part II: North America*, edited by: Ehlers, J. and Gibbard, P. L., 373–424, Elsevier, Amsterdam, doi:10.1016/S1571-0866(04)80209-4, 2004.
- Engelhart, S. and Horton, B. P.: Holocene sea level database for the Atlantic coast of the United States, *Quaternary Sci. Rev.*, 54, 12–25, doi:10.1016/j.quascirev.2011.09.013, 2012.
- Engelhart, S., Peltier, W. R., and Horton, B. P.: Holocene relative sea-level changes and glacial isostatic adjustment of the U.S. Atlantic coast, *Geology*, 39, 751–754, doi:10.1130/G31857.1, 2011.
- Engelhart, S. E., Vacchi, M., Horton, B. P., Nelson, A. R., and Kopp, R. E.: A sea-level database for the Pacific coast of central North America, *Quaternary Sci. Rev.*, 113, 78–92, doi:10.1016/j.quascirev.2014.12.001, 2015.
- Godwin, H.: Studies in the post-glacial history of British vegetation. III: Fenland pollen diagrams. IV: Post-glacial changes of relative land and sea level in the English Fenland, *Philos. T. R. Soc. B*, 230, 239–303, doi:10.1098/rstb.1940.0001, 1940.
- Hijma, M., Engelhart, S., Horton, B. P., Törnqvist, T., Hu, P., and Hill, D.: A Protocol for a Geological Sea-Level Database, in: *Handbook of Sea-Level Research*, edited by: Shennan, I.,

- Long, A. J., and Horton, B. P.: Wiley Blackwell, 536–553, doi:10.1002/9781118452547.ch34, 2015.
- Hirsch, J. E.: An index to quantify an individual's scientific research output, *P. Natl. Acad. Sci. USA*, 102, 16569–16572, doi:10.1073/pnas.0507655102, 2005.
- Horton, B. P., Edwards, R. J., and Floyd, J. M.: Implications of a microfossil-based transfer function in Holocene sea-level studies, *Geol. Soc. Sp.*, 166, 41–54, doi:10.1144/GSL.SP2000.166.01.03, 2000.
- Horton, B. P., Gibbard, P. L., Mine, G. M., Morley, R. J., Purintavaragul, C., and Stargardt, J. M.: Holocene sea levels and palaeoenvironments, Malay-Thai Peninsula, southeast Asia, *The Holocene*, 15, 1199–1213, doi:10.1191/0959683605hl891rp, 2005.
- Hughes, A. L. C., Gyllencreutz, R., Lohne, Ø. S., Mangerud, J., and Svendsen, J. I.: The last Eurasian ice sheets – a chronological database and time-slice reconstruction, *DATED-1, Boreas*, 45, 1–45, doi:10.1111/bor.12142, 2016.
- Kattage, J., Diaz, S., and Wirth, C.: Of carrots and sticks, *Nat. Geosci.*, 7, 778–779, doi:10.1038/ngeo2280, 2014.
- Khan, N. S., Ashe, E., Shaw, T. A., Vacchi, M., Walker, J., Peltier, W., Kopp, R. E., and Horton, B. P.: Holocene Relative Sea-Level Changes from Near-, Intermediate-, and Far-Field Locations, *Curr. Climate Change Reports*, 1, 247–262, doi:10.1007/s40641-015-0029-z, 2015.
- Klemann, V., Unger, A., Hibbert, F., and Dransch, D.: SLIVISU, a concept of an interactive visualization framework for the analysis of geological data, *EGU General Assembly Conference Abstracts*, 15, EGU2013–13953, 2013.
- Kopp, R. E., Simons, F. J., Mitrovica, J. X., Maloof, A. C., and Oppenheimer, M.: Probabilistic assessment of sea level during the last interglacial stage, *Nature*, 462, 863–867, doi:10.1038/nature08686, 2009.
- Kopp, R. E., Kemp, A. C., Bittermann, K., Horton, B. P., Donnelly, J. P., Gehrels, W. R., Hay, C. C., Mitrovica, J. X., Morrow, E. D., and Rahmstorf, S.: Temperature-driven global sea-level variability in the Common Era, *P. Natl. Acad. Sci. USA*, 113, E1434–E1441, doi:10.1073/pnas.1517056113, 2016.
- Long, A. J., Woodroffe, S. A., Roberts, D. H., and Dawson, S.: Isolation basins, sea-level changes and the Holocene history of the Greenland Ice Sheet, *Quaternary Sci. Rev.*, 30, 3748–3768, doi:10.1016/j.quascirev.2011.10.013, 2011.
- Milne, G. A., Long, A. J., and Bassett, S. E.: Modelling Holocene Relative Sea-Level Observations from the Caribbean and South America, *Quaternary Sci. Rev.*, 24, 1183–1202, doi:10.1016/j.quascirev.2004.10.005, 2005.
- Muhs, D. R., Simmons, K. R., Schumann, R. R., and Halley, R. B.: Sea-level history of the past two interglacial periods: new evidence from U-series dating of reef corals from south Florida, *Quaternary Sci. Rev.*, 30, 570–590, doi:10.1016/j.quascirev.2010.12.019, 2011.
- Overpeck, J. T., Meehl, G. A., Bony, S., and Easterling, D. R.: Climate data challenges in the 21st century, *Science*, 331, 700–702, doi:10.1126/science.1197869, 2011.
- Paskin, N.: Digital Object Identifiers for scientific data, *Data Sci. J.*, 4, 12–20, doi:10.2481/dsj.4.12, 2005.
- Pedoja, K., Husson, L., Johnson, M. E., Melnick, D., Witt, C., Pochat, S., Nexer, M., Delcaillau, B., Pinegina, T., Poprawski, Y., Authemayou, C., Elliot, M., Regard, V., and Garestier, F.: Coastal staircase sequences reflecting sea-level oscillations and tectonic uplift during the Quaternary and Neogene, *Earth-Sci. Rev.*, 132, 13–38, doi:10.1016/j.earscirev.2014.01.007, 2014.
- Peltier, W. R., Argus, D. F., and Drummond, R.: Space geodesy constrains ice age terminal deglaciation: The global ICE-6G\_C (VM5a) model, *J. Geophys. Res.-Sol. Ea.*, 120, 450–487, doi:10.1002/2014JB011176, 2015.
- Prest, V., Grant, D. R., and Rampton, V. N.: Glacial Map of Canada, Map 1532A, Ottawa, doi:10.4095/108979, 1968.
- Quadt, F., Düsterhus, A., Höck, H., Lautenschlager, M., Hense, A. V., Hense, A. N., and Dames, M.: Atarrabi – A workflow system for the Publication of Environmental Data, *Data Sci. J.*, 11, 89–109, doi:10.2481/dsj.012-027, 2012.
- Reynolds, L. C. and Simms, A. R.: Late Quaternary relative sea level in Southern California and Monterey Bay, *Quaternary Sci. Rev.*, 126, 57–66, doi:10.1016/j.quascirev.2015.08.003, 2015.
- Rovere, A., Raymo, M. E., O'Leary, M. J., and Hearty, P. J.: Crowdsourcing in the Quaternary sea level community: insights from the Pliocene, *Quaternary Sci. Rev.*, 56, 164–166, doi:10.1016/j.quascirev.2012.09.014, 2012.
- Rovere, A., Raymo, M. E., Mitrovica, J. X., Hearty, P. J., O'Leary, M. J., and Inglis, J. D.: The Mid-Pliocene sea-level conundrum: Glacial isostasy, eustasy and dynamic topography, *Earth Planet. Sc. Lett.*, 387, 27–33, doi:10.1016/j.epsl.2013.10.030, 2014.
- Rovere, A., Hearty, P. J., Austermann, J., Mitrovica, J. X., Gale, J., Moucha, R., Forte, A. M., and Raymo, M. E.: Mid-Pliocene shorelines of the US Atlantic Coastal Plain – An improved elevation database with comparison to Earth model predictions, *Earth-Sci. Rev.*, 145, 117–131, doi:10.1016/j.earscirev.2015.02.007, 2015.
- Rowley, D. B., Forte, A. M., Moucha, R., Mitrovica, J. X., Simons, N. A., and Grand, S. P.: Dynamic Topography Change of the Eastern United States Since 3 Million Years Ago, *Science*, 340, 1560–1563, doi:10.1126/science.1229180, 2013.
- Roy, K. and Peltier, W. R.: Glacial isostatic adjustment, relative sea level history and mantle viscosity: reconciling relative sea level model predictions for the U.S. East coast with geological constraints, *Geophys. J. Int.*, 201, 1156–1181, doi:10.1093/gji/ggv066, 2015.
- Shakun, J. D., Clark, P. U., He, F., Lifton, N. A., Liu, Z., and Otto-Bliesner, B. L.: Regional and global forcing of glacier retreat during the last deglaciation, *Nat. Comm.*, 6, 8059, doi:10.1038/ncomms9059, 2015.
- Shennan, I.: Flandrian sea-level changes in the Fenland. II: Tendencies of sea-level movement, altitudinal changes, and local and regional factors, *J. Quaternary Sci.*, 1, 155–179, doi:10.1002/jqs.3390010205, 1986.
- Shennan, I. and Horton, B.: Holocene land- and sea-level changes in Great Britain, *J. Quaternary Sci.*, 17, 511–526, doi:10.1002/jqs.710, 2002.
- Shennan, I., Long, A. J., and Horton, B. P. (Eds.): *Handbook of Sea-Level Research*, Wiley Blackwell, doi:10.1002/9781118452547, 2015.
- Siddall, M., Abe-Ouchi, A., Andersen, A., Antonioli, F., Bamber, J. L., Bard, E., Clark, J., Clark, P. U., Deschamps, P., Dutton, A., Elliot, M., Gallup, C., Gomez, N., Gregory, J. M., Huybers, P., Kawarnura, K., Kelly, M., Lambeck, K., Lowell, T., Mitrovica, J. X., Otto-Bliesner, B. L., Richards, D., Stanford, J. R., Stirling, C., Stocker, T. F., Thomas, A. L., Thompson, B., Törnqvist, T.,

- Riveiros, N., Waelbroeck, C., Yokoyama, Y., and Yu, S. U.: The sea-level conundrum: case studies from palaeo-archives, *J. Quaternary Sci.*, 25, 19–25, doi:10.1002/jqs.1270, 2010.
- Stokes, C. R., Tarasov, L., Blomdin, R., Cronin, T. M., Fisher, T. G., Gyllencreutz, R., Hättestrand, C., Heyman, J., Hindmarsh, R. C., Hughes, A. L., Jakobsson, M., Kirchner, N., Livingstone, S. J., Margold, M., Murton, J. B., Noormets, R., Peltier, W. R., Peteet, D. M., Piper, D. J., Preusser, F., Renssen, H., Roberts, D. H., Roche, D. M., Saint-Ange, F., Stroeve, A. P., and Tellert, J. T.: On the reconstruction of palaeo-ice sheets: Recent advances and future challenges, *Quaternary Sci. Rev.*, 125, 15–49, doi:10.1016/j.quascirev.2015.07.016, 2015.
- Stroeve, A. P., Hättestrand, C., Kleman, J., Heyman, J., Fabel, D., Fredin, O., Goodfellow, B. W., Harbor, J. M., Jansen, J. D., Olsen, L., Caffee, M. W., Fink, D., Lundqvist, J., Rosqvist, G. C., Strömberg, B., and Jansson, K. N.: Deglaciation of Fennoscandia, *Quaternary Sci. Rev.*, doi:10.1016/j.quascirev.2015.09.016, online first, 2015.
- Tarasov, L., Dyke, A., Neal, R. M., and Peltier, W. R.: A data-calibrated distribution of deglacial chronologies for North American ice complex from glaciological modelling, *Earth Planet. Sc. Lett.*, 315–316, doi:10.1016/j.epsl.2011.09.010, 2012.
- Törnqvist, T. E., Rosenheim, B. E., Hu, P., and Fernandez, A. B.: Radiocarbon dating and calibration, in: *Handbook of Sea-Level Research*, edited by: Shennan, I., Long, A. J., and Horton, B. P., 349–360, Wiley Blackwell, doi:10.1002/9781118452547.ch23, 2015.
- Unger, A., Schulte, S., Klemann, V., and Dransch, D.: Visual Analytics Concept for the Validation of Geoscientific Simulation Models, *IEEE Trans. Vis. Comp. Graph.*, 18, 2216–2225, doi:10.1109/tvcg.2012.190, 2012.
- Vacchi, M., Rovere, A., Chatzipetros, A., Zouros, N., and Firpo, M.: An updated database of Holocene relative sea level changes in NE Aegean Sea, *Quaternary Int.*, 328–329, 301–310, doi:10.1016/j.quaint.2013.08.036, 2014.
- Vacchi, M., Marriner, N., Morhange, C., Spada, G., Fontana, A., and Rovere, A.: Multiproxy assessment of Holocene relative sea-level changes in the western Mediterranean: Sea-level variability and improvements in the definition of the isostatic signal, *Earth-Sci. Rev.*, 155, 172–197, doi:10.1016/j.earscirev.2016.02.002, 2016.
- van de Plassche, O.: *Sea-level research: a manual for the collection and evaluation of data*, Springer Netherlands, doi:10.1007/978-94-009-4215-8, 1986.
- Vaughan, D. G., Comiso, J., Allison, I., Carrasco, J., Kaser, G., Kwok, R., Mote, P., Murray, T., Paul, F., Ren, J., Rignot, E., Solomina, O., Steffen, K., and Zhang, T.: Observations: Cryosphere, in: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P., Cambridge University Press, Cambridge, UK, and New York, NY, USA, 2013.
- Vink, A., Steffen, H., Reinhardt, L., and Kaufmann, G.: Holocene relative sea-level change, isostatic subsidence and the radial viscosity structure of the mantle of northwest Europe (Belgium, the Netherlands, Germany, southern North Sea), *Quaternary Sci. Rev.*, 26, 3249–3275, doi:10.1016/j.quascirev.2007.07.014, 2007.
- Wahl, E. R., Anderson, D. M., Bauer, B. A., Buckner, R., Gille, E. P., Gross, W. S., Hartman, M., and Shah, A.: An archive of high-resolution temperature reconstructions over the past 2+millennia, *Geochim. Geophys. Geosy.*, 11, doi:10.1029/2009GC002817, 2010.
- Whitehouse, P. L., Bentley, M. J., and Le Brocq, A. M.: A deglacial model for Antarctica: geological constraints and glaciological modelling as a basis for a new model of Antarctic glacial isostatic adjustment, *Quaternary Sci. Rev.*, 32, 1–24, doi:10.1016/j.quascirev.2011.11.016, 2012.
- Zong, Y.: Mid-Holocene sea-level highstand along the Southeast Coast of China, *Quaternary Int.*, 117, 55–67, doi:10.1016/S1040-6182(03)00116-2, 2004.