Dear David,

Please find a revised version of our manuscript: “Integrating palaeoclimate time series with rich metadata for uncertainty modelling: strategy and documentation of the PALMOD 130k marine palaeoclimate data synthesis”. We would like to thank the reviewers for their helpful feedback on the previous version. We have addressed all their comments and a detailed response, as well as an updated version of the manuscript with tracked changes, is provided below. We feel that the manuscript is now in a better shape and hope that it meets the criteria for publication in Earth System Science Data.

In addition to changes to the text we have also improved the standardisation and corrected some minor mistakes (e.g. typos) in the metadata. We have not added any sites, nor added data, so the figures in the manuscript are still up to date. The updated data product (version 1.0.1; https://seafile.zfn.unibremen.de/d/c146404ef5634c2497c2/) has been submitted for long-term archiving at PANGAEA. However, due to the current corona crisis, PANGAEA is at the moment receiving more submission whilst at the same time dealing with personnel shortages, which means that the processing time can be up to several weeks. We hope that we can provide a new download link to the data soon and any updates to the data product will also be clearly highlighted at the PANGAEA page where version 1.0.0 is located (https://doi.pangaea.de/10.1594/PANGAEA.908831). As said above, this update does not in any way affect the data themselves and the data descriptor is therefore not affected. We therefore hope that the revised manuscript can be considered for publication at this stage.

With kind regards,

Lukas Jonkers
(on behalf of all authors)
We would like to thank the reviewer for their constructive comments. We copied the comments below and provide our proposed changes in red. We hope that these changes address the concerns and that a revised manuscript will meet the criteria for publication in Earth System Science Data.

Lukas Jonkers
(On behalf of all authors)

The authors present a compilation of paleoclimate data from marine sediment cores covering the past 130 kyr. They give a clear account of their data acquisition strategy, which focussed on cores with d18O measured on benthic foraminifera and radiocarbon dates so that a robust common chronology could be constructed for the entire compilation. Where other paleoclimate data were available for the same core, these data were included in the compilation. They pay close attention to including meta data required to analyse the data further.

Parameter and metadata names have been harmonised and the original naming is preserved so that these can be traced in the original publications. The data are all well referenced with DOIs and citations. New depth-age models have been constructed for all sites using BACON and the published chronologies preserved for reference.

In all this represents a very well researched and harmonised dataset with rich and useful metadata that does not exist elsewhere. The data are supplied in a variety of formats, as R data objects, NetCDF, and in the LiPD format which is itself a set of zipped plain text (csv) files containing the data in a highly structured JSON format (http://wiki.linked.earth/Linked_Paleo_Data). We appreciate the reviewer’s feedback on the value of the data product.

However, the structure of the data within the R objects makes it very difficult to search for and extract subsets of data. For example, all the records of variable “planktonic.d18O”, or all the records in a certain geographical region. In the NetCDF and LiPD formats the data are also structured in a similar way, although there may be tools available to help work with the LiPD data.

Use of this data compilation would be greatly enhanced if the data were restructured into a set of “partially normalized tables” in a “star schema” so that queries can be made in an SQL-like way by joining tables and using select and filter type statements. See Brian McGill’s 3rd Commandment here (https://dynamicecology.wordpress.com/2016/08/22/ten-commandments-for-good-data-management/). A key table in this format would for example be the “ParameterListWithRefs.csv” table linked to in the Data availability statement of this manuscript but not found within the data objects.

No specific database software needs to be used; these could be plain text files that could be read in by many data analysis software. This is not a “big” dataset so the structure does not need to optimise storage or retrieval efficiency.

I’m not suggesting that reformatting the data in this way would be trivial for the authors, but the data in their current format are well structured and so it should be possible to write code to do it – and this should be much easier for the authors than for someone coming to it fresh.

We thank the reviewer for their feedback on the format of the database. Because of its structure, linking multiple records to common site information, complex metadata and chronological information, the data we provide are indeed not as easy to query as a simple text file would be. Yet, the referee is right to point out that the ability to extract user-defined sets of records, such as by regions or by proxy type, is a key functionality. We have therefore provided example scripts to query the RDS files and made them available on GitHub (https://github.com/lukasjonkers/PALMODutils). We would like to highlight that similar tools are already available to query LiPD files (https://github.com/nickmckay/LiPD-utilities) and that the PaleoDataView software is specifically designed to interact with the netcdf files (https://www.marum.de/en/Stefan-Multitz/PaleoDataView.html). Considering that for both the LiPD and...
netCDF formats querying tools are available, and that we provide similar options for R, we prefer to keep the format of the database as it is now, without shoehorning its highly interlinked structure into another format. This is because the structure of the database also allows for better quality control and updating of the (meta)data and age-depth models. The individual files also make unintentional mixing of information less likely.

The example scripts that we will provide will allow the user to build their own version of a table like ‘ParameterListWithRefs.csv’ and to query the database by:

- Parameter
- Parameter detail
- Sensor species
- Minimum age
- Maximum age
- Resolution
- Number of tie points

The example code returns the indices of the sites that meet the criteria, which the user can use to extract the desired time series data, metadata or chronology data. This is intentional as custom scripts are required to tailor the data extraction to each analysis. We have updated section 6 and incorporated the information on how to interact with the database.

Minor comment: l. 486 - In the section “recommendations for data archiving” “Include metadata” I would also recommend including information about the size of the sample on which the parameter was measured, e.g. number of foraminifera, mass of sample, total peak area for Alkenones. As this can be very useful when assessing the uncertainty of the value.

We agree and will include a statement encouraging researchers to include those data in the section with recommendations. Note that this also follows the recommendation of PaCTS (Khider et al., 2019).

Text errors:
l. 147 “were” -> “where”
l. 161 “more of data”
These will be addressed.

References:

We would like to thank Blanca Ausin for her constructive comments. We copied the comments below and provide our proposed changes in red. We hope that these changes address the concerns and that a revised manuscript will meet the criteria for publication in Earth System Science Data.

Lukas Jonkers
(On behalf of all authors)
The authors present a database consisting of paleoclimate records from several climate-sensitive parameters from 143 marine sites spanning the last 130 kyr. New chronologies have been built for each site for consistency while corresponding original chronologies are also provided. One advantage over other data compilations is the careful treatment of chronological data. The database contains rich metadata needed to perform a robust assessment of the paleoclimate signals and is available in three different formats facilitating its use in free software. The manuscript is clearly written and has a straightforward explanation of the data search and treatment strategy, database structure, data storage and future plans.

The authors provide good rationale on the need of a database of time series data for modelling purposes and the importance of data formatting and standardization and their effort to build the presented database should be welcomed by the paleoclimate and climate modeler community. There are a couple of comments I raise below I think the authors should take care of:

Lines 257-265: The use of variable reservoir ages to build age-depth models is a hot topic that is currently under debate. So far, no clear consensus exists on whether the general use of this or the “static reservoir” approach is more robust, advantageous, or beneficial than the other and the choice is made by every author based on different reasons. A discussion on why the authors decide to apply this approach should be included.

The reviewer rightly points out the non-trivial issue of reservoir corrections in age-depth modelling. It has been established beyond doubt that reservoir ages vary in space and time, so the assumption of a static reservoir age is not warranted. Its attractiveness is only in its simplicity as it allows for straightforward age-depth modelling, especially given the uncertainty of how reservoir ages varied in the past. Therefore, we feel that our approach is more realistic as it attempts to incorporate the variability of reservoir ages. Because it is based on a model approach, it has the additional advantage that it does so in a globally physically constrained way. Our original wording also reflects this: “To ensure a common chronological framework for all time series in the synthesis, radiocarbon ages were re-calibrated using the IntCal13 curve (Reimer et al., 2013). We used reservoir age estimates (including uncertainty) based on a comprehensive ocean general circulation model (Butzin et al., 2017) to account for physically plausible spatial and temporal variability in the reservoir ages.” That said, it is of course not given that the model by Butzin et al. (2017) is the ultimate answer to the question of how reservoir ages varied in the past. This work will be superseded by new studies and indeed in many situations, like for comparison with published work, the user may wish to see age-depth models calibrated to constant reservoir ages of different kinds. We have anticipated this development, which is why all information on age-depth model development is kept in the database and can be used to derive alternative models (see e.g. section 2.3). We felt it important to provide uniformly derived age estimates for all records and believe the approach we use is not inferior to any existing alternative. What we would like to avoid is to overinflate the dataset by providing multiple age-depth models for each site. In the revised manuscript we will better explain our choice to use variable reservoir ages, but as we provide the necessary data for the user to amend the age-depth models we think that a discussion about which approach is more robust is moot/does not add to the manuscript. Given the comment below, we have however also improved the description of how we obtained the reservoir ages and assessed their uncertainty (section 2.3).

Also, I find their specific approach strongly relies on Butzin et al. (2017) model and is limited by the spatial coverage of the latter. How did the authors decide on which reservoir ages to use for sediment cores located out of the original data coverage of Butzin et al. (2017)? What is the uncertainty associated with the reservoir ages taken from the extrapolated regions in Butzin et al. (2017) and do the authors account for it? More specifically, do the authors account for additional uncertainty of reservoir ages for cores retrieved from upwelling regions and semi-isolated basins (if any is included in the database) where the effect of regional processes are not considered by the coarse-resolution of most general circulation models?

The reviewer is correct that our approach to take a variable reservoir age into account is limited by the resolution (3.5°) of the data set from Butzin et al. (2017). In most current studies, marine radiocarbon age models are based on a calibration with the Marine13 calibration curve, which is already corrected by a
global modelled reservoir age and can be further corrected with a local dR that is usually assumed to be constant. However, although the Butzin et al. (2017) data set can only be a first-order approximation, we see several advantages in its application: First the dataset allows us to take some reasonable regional variation of the reservoir age into account. For example, the high glacial reservoir ages in the North Atlantic (i.e. Sarnthein et al. 2015, Radiocarbon, 57(1), 129–151, doi:10.2458/azu_rc.57.17916) are well reproduced by Butzin et al. (2017). Second, Butzin et al. (2017) consider the effect of pCO2 changes on the global reservoir age of the surface ocean that may lead to a glacial global reservoir age of 600-700 years compared to a global reservoir age of 405 years taken into account for most of Marine13. To derive the reservoir age/reservoir age error for a measured radiocarbon age we (i) extract all modelled radiocarbon ages from the nearest gridbox in the modelled data set, (ii) find all modelled radiocarbon ages that are possible within the error of the measured radiocarbon age and then (iii) take the mean and the standard deviation of all corresponding reservoir ages as correction for the measured radiocarbon age. Due to a lack of information, we cannot account for local uncertainties in the reservoir ages, i.e. due to upwelling. To clarify, we have added further information, how reservoir ages have been derived (section 2.3). The issue of reservoir ages in semi-enclosed basins such as the Mediterranean and Red Seas, was already described in the original manuscript in section 2.3.

SST is a vital component of a vast majority of modelling efforts. I highly encourage the authors to recalibrate SST estimates from alkenone ratios with the latest global calibration by Tierney and Tingley (2018). This does not differ much from previous calibrations for estimates below 24°C, but over that value it corrects for the slope attenuation in the UK’37-SST relationship providing temperature estimates several degrees different from previous estimates. This calibration also provides error estimates and specific calibrations for regions where seasonality has a strong influence of alkenones-derived SST. We agree with the reviewer that calibration of temperature proxies is an important issue. Precisely for this reason we have included the raw data in the database whenever this was possible to allow the user of the database to recalibrate any temperature proxy using the calibration they want/deem most appropriate. At this stage we refrain from recalibrating UK37 temperature estimates because we plan updates of the database that focus more on temperature. In these updates we will ensure globally consistent calibration of all temperature proxies. In addition, we feel that changing the original calibration would require more consideration than just changing the calibration equation, as the calibration choice of the authors of each individual dataset are likely to have clear reasons for using each specific calibration. We therefore think that recalibrating all temperature proxies goes beyond the scope of the present manuscript. Recalibrating only the UK37 temperatures seems in our opinion arbitrary given that progress has also been made on Mg/Ca temperatures, transfer functions etc. We stress however, that any user of the database can with relatively little effort change the calibration that was used by the authors of the data, as long as raw data were made available.

Minor comments:
Line 28: Why there are more benthic δ18O time series than sites?
This is a good question. It is because some sites contain data from more than a single species, or multiple time-series from the same species that could not be merged, for instance because of differences in size fraction of the shells used for the measurements.
Line 96: What do you mean by homogenous? Please add continuous chronology.
Clarified. The new sentence reads: “We focus on the ocean as it is a large reservoir of heat and CO2 and allows global coverage with consistent chronological control.” L: 96.
Lines 122 and 123: the term “marine sediment sequences” might be more appropriate than “marine sediment archives”.
We prefer to keep the original wording.
Line 132: Please state somewhere in this paragraph why the need for combining radiocarbon and benthic δ18O for the chronological control (e.g., radiocarbon limit).
Good point. We inserted the following sentence: “This approach of blending absolute and relative age controls is required to provide age-depth models for sediment cores that extend beyond the radiocarbon dating range (~40,000 years).” L:10-142.
Line 151-156: why so self-critical? Chronostratigraphies and age-depth models are the backbone of paleoclimate data. I find the choice of stating a robust and consistent chronological control as the conditioning criteria to build the stratigraphic framework an advantage rather than a pitfall, even if the database is not as comprehensive as it could be by using more flexible criteria for the time control.
We agree with the reviewer and are happy with their support for our approach. However, we would like to stick with the current wording because we envision updates of the database that will include different age modelling strategies.
Line 161: rephrase “by including more data”.
Done.
Line 181: rephrase “we followed the reduction approach”.
Done.
Line 218: Please mention tables in sequential order.
Done.
Line 233: in order to
Done.
Line 310: those cases when it is not possible to indicate the calibration are because this is not reported in the original paper?
Yes, we have made this clearer.
Line 312: Species and number of individuals are reported as this is important information to assess foraminifera-derived proxy data. The same is true for foraminifera size, which indeed is reported in the database. Please add here the effect of the size of foraminiferal tests on derived measurements.
Done.
Line 321: The Uk’37 ratio is not based on alkenones with different chain lengths (both have 37 carbons) but with a different number of unsaturations (di- and tri-unsaturated). Please correct.
Silly mistake. Thank you for pointing this out. We have corrected this.
Line 323: Please replace: not constant by “seasonal”.
We prefer our original wording because seasonal suggests that no production occurs outside a specific season, whereas all we want to say is that the abundance of foraminifera is variable over the course of a year - such variability may follow a different pattern than strictly seasonal.
Line 326: C37 alkenone concentration is included in the database but its significance as productivity proxy is not included in the discussion. Is there any reason for this?
The reason to include concentration data was that these data may help to assess the reliability of the temperature estimates, which is now also mentioned in L543-546. Alkenone concentration data can only be assumed to reflect productivity where the overall productivity scales linearly with coccolithophore production.
Line 364: Replace Table by table.
Done.
Line 383: Replace table by Table.
Done.
Line 428: Does this mean no more 14C-based chronologies? Why?
Yes, this is correct. To make clear why we do this we changed the sentence to “To increase the spatio-temporal coverage over the entire 130,000-year time frame of the database, updates of this data product will first aim for quantitative growth of the database by adding more time series with chronological control based on benthic foraminifera δ-18O and absolute age control points other than 14C.”
Line 430: Is there any plan for authors submitting their data for inclusion in the database?
The reviewer raises an important point. Our database is not intended as a replacement of data repositories and we only include data that is publicly available and citable to encourage good data stewardship. This is why we have designed the data synthesis workflow in such a way that for updates of the database data files from PANGAEA and NCDC can be ingested in a straightforward manner. Such updates would benefit hugely from the authors depositing their data in a way that follows the recommendations as spelled out in our manuscript. This applies in particular to the inclusion of metadata and chronological information, which is the main point of our section on data archiving. In this regard we would also like to highlight that NCDC now accepts data in LiPD format, which greatly simplifies the inclusions of metadata and standardisation.
Line 470: The observation time series?
We deleted “observation”.

Line 487: Based on planktonic foraminifera Mg/Ca. . . Please rephrase.
We deleted “Mg/Ca”.

Line 513: Replace Pangaea by PANGAEA
Done.

Line 518: . . . data consists data not part of publication. Rephrase.
Rephrased.

Table 4: it might be convenient to replace “detritus” by “debris” for searching purposes as the latter is more common in the paleoclimate field.

Ice-rafted detritus results in more hits in Google Scholar, so we prefer to keep this naming. We deem the chance of confusion minimal.
Integrating palaeoclimate time series with rich metadata for uncertainty modelling: strategy and documentation of the PALMOD 130k marine palaeoclimate data synthesis

Lukas Jonkers1*, Olivier Cartapanis2, Michael Langner1, Nick McKay3, Stefan Mulitza1, Anne Strack1, Michal Kucera1

* corresponding author: ljonkers@marum.de

1: MARUM - center for marine environmental sciences. University of Bremen, Bremen, Germany.
2: Institute of Geological Sciences and Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland.
3: School of Earth & Sustainability, Northern Arizona University, Flagstaff, USA.

Abstract

Palaeoclimate data hold the unique promise of providing a long-term perspective on climate change and as such can serve as an important benchmark for climate models. However, palaeoclimate data have generally been archived with insufficient standardisation and metadata to allow for transparent and consistent uncertainty assessment in an automated way. Thanks to improved computation capacity, transient palaeoclimate simulations are now possible, calling for data products containing multi-parameter time series rather than information on a single parameter for a single time slice. Efforts are underway to simulate a complete glacial-interglacial cycle using general circulation models (palmod.de) and to confront these simulations with palaeoclimate data, we have compiled a multi-parameter marine palaeoclimate data synthesis that contains time series spanning 0 to 130,000 years ago. We present the first version of the data product that focuses exclusively on time series for which a robust chronology based on benthic foraminifera δ¹⁸O and radiocarbon dating is available. The product contains 896 time series of eight palaeoclimate parameters from 143 individual sites, each associated with rich metadata, age-depth model ensembles and information to refine and update the chronologies. This version contains 205 time series of benthic foraminifera δ¹⁸O, 169 of benthic foraminifera δ¹³C, 131 of seawater temperature, 174 and 119 of planktonic foraminifera δ¹⁸O and δ¹³C and 44, 38 and 16 of carbonate, organic carbon and biogenic silica content, respectively. The data product is available in three formats (R, LiPD and netCDF) facilitating use across different software and operating systems and can be downloaded at https://doi.pangaea.de/10.1594/PANGAEA.908831 (Jonkers et al., 2019). This data descriptor presents our data synthesis strategy and describes the contents and format of the data product in detail. It ends with a set of recommendations for data archiving.
1. Introduction

Global climate has varied dramatically over the last glacial-interglacial cycle. Since the previous interglacial (approximately 130,000 years ago) the earth has slowly cooled until the Last Glacial Maximum (LGM; approximately 21,000 years ago). This cooling was associated with the growth of massive ice sheets in North America and Eurasia, leading to a sea level drop of about 120 metres (Waelbroeck et al., 2002) and pronounced climate variability on millennial time scales (Voelker and et al., 2002). Since the LGM, the Earth has warmed rapidly until the onset of the current relatively stable warm period, the Holocene (Shakun et al., 2012). The ultimate cause of the large-scale variations in the Earth’s climate are changes in the orbit of the Earth around the Sun (Hays et al., 1976). However, complex feedback and non-linear mechanisms, involving ocean (atmosphere, cryosphere) circulation and biogeochemical cycles, are required to explain how slow changes in the orbital configuration led to the observed evolution of global climate, and how these processes led to the manifestation of abrupt climate change.

For these reasons the last glacial-interglacial cycle has been a key target for palaeoclimate modelling. Initially this only involved equilibrium simulations for key time slices, such as the LGM, or transient simulations for short periods, such as the last millennium. The motivation to simulate past climate states is given by the possibility for palaeoclimate data to serve as a benchmark for the models. Indeed, this possibility contributed to the development of large palaeodata syntheses (CLIMAP project members, 1981; MARGO project, 2009). The time-slice modelling approach is still being pursued, for example in phase 4 of the Paleoclimate Intercomparison Project (PMIP), four of the five target intervals fall into the time frame of the last glacial cycle (Kageyama et al., 2018). However, with increasing computing power, the focus is now shifting towards transient climate simulations (Liu et al., 2009; Latif et al., 2016) and the simulation of the last deglaciation is now also considered in the PMIP protocol (Ivanovic et al., 2016).

This development calls for a different type of palaeodata synthesis, with focus on time series rather than time slices. Time series of climate data are needed to evaluate aspects of transient simulations that are not available in equilibrium simulations, such as rates of change, phase relationships and spectral properties of climate variability. It is also clear that an evaluation in multi-parameter space using different aspects of the climate system and multiple proxies will be more powerful and diagnostic (Kurahashi-Nakamura et al., 2017), calling for multi-parameter synthesis products.

Observations of the evolution of past climate are based on proxies (measurable approximations of climate-related variables) and hence are, by definition, indirect. Comparison of proxy-based reconstructions with climate model simulations is therefore far from straightforward, as discrepancies may arise from both model and proxy uncertainty. Proxy uncertainty derives from reconstruction uncertainty (related to calibration, recording bias, archive specifics and instrumental approach) and
chronological uncertainty. The latter is particularly relevant for the comparison of transient climate change and chronological uncertainty thus requires a comprehensive treatment in data syntheses of palaeoclimate time series.

Accounting for proxy uncertainties in a comprehensive and transparent manner requires not only expert knowledge, but also the availability of extensive metadata in addition to the proxy data. However, due to a lack of standardisation and inconsistent archiving of metadata, synthesizing palaeoclimate data in a way that allows robust uncertainty assessment remains challenging and time consuming. Efforts are underway to alleviate these challenges. The largest palaeoclimate data repositories (World Data Service for Paleoclimatology, operated by the national centers for environmental information (NCEI) at NOAA and PANGAEA) are both striving for more standardisation and to store data in (more) machine-readable format. In addition, standardisation is progressing through the use of existing data formats from other communities (netCDF; Langner and Mulitza, 2019) as well as the implementation of new data formats specifically targeted to palaeodata (Linked PaleoData (LiPD); McKay and Emile-Geay, 2016). At the same time there is ongoing discussion on data and metadata requirements and standards (Khider et al., 2019). Traceability of datasets is also improved through data citations, ensuring not only that data producers receive proper credit for their work, but also allowing for better linking of different datasets. Nevertheless, these initiatives are only recently emerging and the majority of the palaeoclimate data remains inconsistently formatted, non-standardised and scattered over various data repositories. The need for synthesis products and documentation of potential synthesis approaches is therefore as large as ever.

Here we present the first version of a new multi-proxy marine palaeoclimate data synthesis that covers the past 130,000 years developed within the German climate modelling initiative PALMOD (Latif et al., 2016). We focus on the ocean as it is a large reservoir of heat and CO₂ and allows global coverage with consistent chronological control. This synthesis goes beyond the time frame of many existing multi-proxy/parameter data syntheses (PAGES2k Consortium et al., 2017; Routson et al., 2019), expands existing data products that provide long palaeoclimatic time series to multiple parameters (Shakun et al., 2012; Marcott et al., 2013; Peterson and Lisiecki, 2018; Snyder, 2016) and is based on a strategy of semi-automated data harvesting (Cartapanis et al., 2016). This version of the synthesis contains data on nine climate-sensitive parameters: benthic and planktonic foraminifera stable oxygen and carbon isotopes, seawater temperature, radiocarbon and bulk sediment carbonate, organic carbon and biogenic silica content.

In this manuscript we describe our synthesis approach, the contents and structure of version 1.0 of the data product, suggestions for its use, plans for future updates and recommendations for archiving new
data and retrieving dark data in a way allowing optimal future re-use. The data product is intended to be used to investigate spatio-temporal changes in a multi-parameter domain. Thanks to rich metadata that allow to rigorously quantify reconstruction uncertainties, we also envision that this data product will provide the building blocks for intelligent palaeoclimate data model comparison (Weitzel et al., 2019), for instance through proxy system modelling (Dolman and Laepple, 2018) or data assimilation (Breitkreuz et al., 2019).

The structure of this data descriptor is as follows. Section 2 describes the synthesis strategy, including the data discovery approach, standardisation and age modelling. In section 3 we provide general information on palaeoclimate proxies from marine sediment archives that is used to guide the metadata selection. Section 4 details the structure of the database and the contents of version 1.0 are outlined in section 5. The formats of the data product and where it can be accessed are described in section 6 and we discuss future plans, versioning and intended use in section 7. In the last section, 8, we reflect on the data synthesis effort and provide recommendations for data archiving and data rescue.

2. Data synthesis strategy

Our data product focuses on time series from marine sediment archives. A single marine sediment archive (sediment core) can be used for measurements of different parameters, each providing information on different aspects of the environmental conditions at the time of deposition. However, for the purpose of analysis, the various proxy time series must refer to a single age-depth model for the sediment core they are derived from. For this reason, the basis of our synthesis is formed by a collection of sediment cores, each associated with its own age-depth model.

Marine sediments are dated using absolute age controls, where specific layers are dated using, for instance, radiocarbon, tephra or palaeomagnetic properties, and/or relative age controls, where time series are aligned based on hypothesised synchronicity of the changes recorded by some properties of the sediment. A well-established hypothesis-based age modelling approach with a solid theoretical basis is the alignment of benthic foraminifera stable oxygen isotope ratio (δ^{18}O) time series (Lisiecki and Raymo, 2005). We thus base our chronological framework on a combination of radiocarbon dates and benthic foraminifera δ^{18}O and have selected time series where both parameters are available as the foundation of this data product. This approach of blending absolute and relative age controls is required to provide age-depth models for sediment cores that extend beyond the radiocarbon dating range (~40,000 years). If available, further proxy time series were then added, thus ensuring a common chronology among all proxy time series measured on the same sediment core.
We selected palaeoclimate parameters to synthesise following discussion with climate modellers within the PALMOD project. The high priority selection includes both physically and biogeochemically relevant parameters, of which some are based on measurements that can be compared with climate model output using (forward proxy) models (e.g. benthic δ¹⁸O) and others represent inferred parameters that can be compared with model output more directly, but for which proxy models are still in their infancy (e.g. temperature based on foraminifera Mg/Ca). Parameter selection also considered the expected spatial and temporal coverage of data availability as well as the existence of previous data products. The high priority parameters for which data are presented here are listed in Table 1. If available, raw data were synthesised and in cases where raw data were not available and it was possible to derive the raw data from the inferred palaeoclimate data, raw data were back-calculated. Raw data time series obtained in this way are flagged with a note describing the calculation.

We note that our approach of first building the stratigraphic framework based on radiocarbon dates and benthic foraminifera δ¹⁸O means that the synthesis is not necessarily comprehensive as it does not include time series where one of the parameters of interest has been measured, but for which the components of the stratigraphic framework are not available. However, at this stage, we opted to include only sediment cores where an age modelling strategy that is consistent and comparable across the entire data product could be achieved.

2.1 Data discovery

In principle, data synthesis can proceed by expansion or reduction (Fig. 1). The first, more traditional, approach relies on expert knowledge of what data is available and/or on systematic literature search. In this approach the synthesis grows by including more data until sufficient data that meet inclusion criteria are compiled. In this way, a lot of time is spent on discovering and retrieving datasets, and it is possible that valuable, but less exposed data is missed. On the other hand, this approach has a chance of uncovering dark data that is not publicly available (Fig. 1).

The second approach starts from a large and “crude” synthesis of data from public sources and proceeds by weeding out data that do not meet criteria for inclusion in the data product. This approach faces different challenges: making sure that the initial bulk database is comprehensive (efficient data mining) and assuring that the data filtering is efficient (fast and accurate). In contrast to the expansion approach, this reduction approach cannot discover dark data. However, it is more objective (less reliant on expert knowledge), can be automated more easily and focuses on data that is already in the public domain so that no time is lost on finding data that ultimately proves unavailable. This second approach also rewards and encourages good data stewardship.
In theory, both approaches can lead to a similarly sized/exhaustive synthesis, but they differ in the allocation of effort (Fig. 1). In practice both approaches are often combined, especially towards the end of a synthesis project, when the data product is benchmarked against existing syntheses.

2.2 Synthesis

2.2.1 Initial synthesis

We followed the reduction approach and used a semi-automated pipeline to compile data from public sources. Keywords (supplementary information) were used to make lists of URLs of potentially relevant data on pangaea.de and the linked files were then downloaded in bulk ($n = 108,239$). A slightly different approach was followed for the NCEI archive. Here, all files that were machine readable at the time of download (September 2016, $n = 1,925$) were obtained from the ftp server (ftp://ftp.ncdc.noaa.gov/pub/data/paleo/paleocean//sediment_files/complete). Custom scripts in R were used to put all data in a common format and merge time series that could be unambiguously assigned to the same core (based on name and $x$, $y$, $z$ position). This resulted in a mixture of records that were merged to the same core and those that could not be either because there was only one data file for the core or because of ambiguous labelling. We refer to the locations of these records as "sites". In order to facilitate the analysis (filtering) of the sites, a uniform attribution of the various parameter names had to be developed. Because no standardised names exist for palaeoclimate parameters, the uniform attribution required the development of attribution libraries for each desired palaeoclimate parameter. The initial synthesis contained time series from 38,511 sites.

2.2.2 Data reduction and standardisation of ontologies

The initial synthesis was reduced by removing non-marine sites (using elevation flag) and further constrained by only considering sites where at least one data point of any of the parameters measured in that core fell within the target time frame (disambiguating age units in the synonym library of the category "age") and the site had benthic oxygen isotope data. This resulted in 781 sites. At this stage, no criteria for length or resolution were applied but we prioritised processing time series that we estimated to contain at least 50 data points within the 130,000-years timeframe. Further data processing started with dereplication of the selected sites. This was necessary because no standards exist for the naming of cores and the repositories store data with different renditions of the same core name, sometimes even associated with erroneous geographic coordinates. This process was done manually and proceeded by constructing a list of disambiguated sites through sequential one-by-one comparison. Where a strict synonym was found (different labels for the same core but the same data), only unique data were retained. At this stage, disambiguation of site names was only done for sites that had at least benthic $\delta^{18}O$, so time series of other parameters, but associated with inconsistent core labels could have been
missed in the synthesis. However, those sites are contained in the initial bulk synthesis and are hence not lost, but will be salvaged in updates of the data product (see section 7).

Further steps required a manual standardisation of the names of the parameters and their attributes (such as species name that was analysed for oxygen isotopes). This was accomplished by deciding on a final, uniform list of parameters and associated metadata and their possible values. Original parameter names were preserved to allow for cross checking. By metadata, we refer to aspects of the individual parameters that were deemed essential to facilitate a meaningful analysis in a palaeoclimatic context, considering potential sources of uncertainty, such as species of foraminifera analysed or calibration equation used for palaeotemperature estimates. The list of metadata is provided in Table 2. The standardisation was accompanied by further dereplication of individual time series that were already associated with the same site name but archived more than once.

2.2.3 Metadata and chronology

Subsequently, as far as possible, metadata values were added manually when missing, often by scraping the information from the original publication. Next, all time series from a single site (core) were put on a common depth scale to allow for age modelling. Data that could not be put on a depth scale were excluded from the synthesis. This was the case where parameter values were recorded only against age and where no other data file was available that allowed unambiguous re-assignment to depth. Ambiguity also resulted from the use of multiple (composite) depth scales for the same archive. Finally, chronological data (all absolute markers, including radiocarbon dates and associated metadata) were manually added, where necessary also by consulting the original publications. Throughout the process, publication information (digital object identifier (DOI), or if not available full bibliographic details) and data source (URL and/or DOI) were preserved in order to trace the source of the data. This applies both to the source of the individual data files from repositories and to the sources of the metadata and chronological data.

2.3 Age modelling

Whereas the initial steps of data discovery and synthesis could rely on published chronology, analysis of the complete dataset will require the development of a common chronological framework. This framework must be constructed in a way that not only allows a consistent method of assignment of ages to depths within each core, but also allows consistent and quantitative assessment of age uncertainty. To this end, we follow an approach that combines absolute ages (radiocarbon ages, tephra layers and palaeomagnetic events) with δ¹⁸O stratigraphy. As a result, our age models may differ from those reported in the original publication(s). This is not to state that the updated age models are better
(constrained), but they are constructed in a way that allows applicability and consistency across the synthesis. The consistent approach allows an assessment of age uncertainty jointly for all records by a Bayesian approach, generating ensembles of sedimentation histories consistent with the available age control points for each core, allowing uncertainty estimates at each depth by considering the distribution of ages given by the ensemble.

With respect to the reporting of the chronology, we follow a transparent approach, preserving the initial age model and providing the new age models as well as all information needed to revise or update the new age models. In the final step, the age information from absolute ages and δ\(^{18}\)O tie-points is combined and the age model and its uncertainty was assessed in a Bayesian framework using BACON (Blaauw and Christen, 2011). The entire age modelling routine was carried out in PaleoDataView (PDV; Langner and Mulitza, 2019).

To ensure a common chronological framework for all time series in the synthesis, radiocarbon ages were re-calibrated using the IntCal13 curve (Reimer et al., 2013). Since reservoir ages vary in space as well as in time, we used reservoir age estimates based on a comprehensive ocean general circulation model (Butzin et al., 2017) to account for this variability in a physically plausible way. To derive the reservoir age and uncertainty for a measured radiocarbon age PDV (i) extracts all modelled radiocarbon ages from the nearest grid cell in the modelled data set, (ii) finds all modelled radiocarbon ages that are possible within the error of the measured radiocarbon age and (iii) takes the mean and the standard deviation of all corresponding reservoir ages as correction for the measured radiocarbon age. By definition, this approach cannot account for processes affecting the reservoir ages on subgrid spatial scales. Given the relatively coarse resolution of the model, this means that processes such as upwelling are not fully accounted for. In addition, no modelled reservoir age data are available for the Mediterranean and Red Seas, we use the reservoir ages reported by the authors of the original publication and an assumed uncertainty of 100 years for these basins (5 sites). Absolute ages based on North Atlantic tephra layers and palaeomagnetic events were updated and harmonised using Svensson et al. (2008).

In addition, and beyond the 14C dating realm (~40,000 years) the age models rely on manual tuning of the benthic foraminifera δ\(^{18}\)O time series from each core to regional benthic foraminifera δ\(^{18}\)O stacks (Lisiecki and Stern, 2016). Stable isotope stratigraphy in theory provides a range of events to correlate, however, in order not to inflate confidence in the tuned age models and ensure comparability between different cores, in our approach, the tuning was done as far as possible only by matching the position of marine isotope stage boundaries. We updated the age-depth models only for the 0-130,000 years time frame of this synthesis, but data and original age models extending beyond 130,000 are preserved in the data product. To obtain uncertainty on the age control points obtained by δ\(^{18}\)O tuning, we used the...
chronological uncertainty of the $\delta^{18}O$ stacks, as reported by Lisiecki and Stern (2016). Additional uncertainty associated with the identification of the control points in the individual records or with the assumption on synchronicity was ignored, as these are difficult, if not impossible, to quantify.

3. Notes on palaeoclimate proxies in marine sediment archives and metadata

This synthesis contains climate sensitive proxy data based on measurements on various biological sensors. It is not the intention here to provide a full overview of marine palaeoclimate proxies and their uncertainties (for this, see e.g. Hillaire-Marcel and De Vernal (2007); Moffa-Sanchez et al. (2019)), but the fact that the proxies are based on biological sensors means that they are affected by different ecological bias in addition to observational noise. Basic knowledge of the recording system is therefore essential for the interpretation of the data and may aid to explain differences between proxies for the same climate parameter. These considerations were also essential to choose the range of metadata to be recorded alongside each palaeoclimatic parameter, to allow a proxy-specific assessment of uncertainty.

Foraminifera are among the most widely used proxy sensors in palaeoceanography. They are unicellular marine zooplankton. The species used here all build a calcite skeleton that is preserved in the sediment. Foraminifera can be divided into two main groups: benthic foraminifera living at the seafloor or at shallow depth in the sediment and planktonic foraminifera living in the upper hundreds of metres of the ocean. In the data product, proxies measured on these two groups are clearly distinguished by a benthic or planktonic prefix. The chemical composition of foraminifera reflects environmental conditions of the seawater the organisms calcified in. For the purpose of this data product the parameters of interest are stable oxygen and carbon isotope and Mg/Ca ratios. Stable oxygen isotope ratios in foraminifera calcite reflect a combination of temperature and $\delta^{18}O$ of seawater (Urey, 1948), which is in turn related to ice volume and salinity. Species-specific calibrations exist to quantitatively link $\delta^{18}O_{\text{foraminifera}}$ and $\delta^{18}O_{\text{seawater}}$ to temperature (e.g. Marchitto et al., 2014; Bemis et al., 1998). Stable carbon isotope ratios ($\delta^{13}C$) reflect the $\delta^{13}C$ of the dissolved inorganic carbon in seawater. In particular the benthic foraminifera species *Cibicidoides wuellerstorffi* generally incorporates $\delta^{13}C_{\text{DIC}}$ without a biological offset and can serve as a tracer of bottom-water $\delta^{13}C_{\text{DIC}}$, which is commonly used as non-passive circulation tracer (Curry and Oppo, 2005). The $\delta^{13}C$ of other benthic foraminifera species is generally not indicative of bottom water $\delta^{13}C_{\text{DIC}}$ and the $\delta^{13}C$ of planktonic foraminifera is also influenced by temperature and carbonate ion concentration, rendering interpretation complicated (Spero et al., 1997).

The Mg/Ca ratio in foraminifera calcite can be used to infer calcification temperature and in combination with $\delta^{18}O_{\text{foraminifera}}$, the $\delta^{18}O_{\text{seawater}}$ (Elderfield and Ganssen, 2000). Similar to stable oxygen isotopes,
Species-specific calibrations exist to quantitatively reconstruct past temperature from Mg/Ca ratios (Anand et al., 2003; Lear et al., 2002) and whenever indicated in the original publication, the calibration is included in the metadata. Carbonate system parameters and salinity have a secondary influence on Mg/Ca ratios in foraminifera calcite (Gray et al., 2018). Whereas benthic foraminifera live in a generally stable environment, the near sea surface habitat of planktonic foraminifera shows large seasonal and vertical gradients. Species-specific seasonal and/or depth habitat preferences may therefore leave a considerable imprint on the proxy signal contained in their shells (Jonkers and Kučera, 2017; Mix, 1987). For all proxies based on foraminifera, it is relevant to record the species as well as the number of individuals that were pooled for geochemical analysis. The latter is because the short life span and variable habitat of foraminifera species causes large variability among individuals. Planktonic foraminifera shell size may for several reason also affect their chemistry (Jonkers et al., 2013; Friedrich et al., 2012) as well as their assemblage composition (Al-Sabouni et al., 2007). Therefore, the size fraction of the analysed shells was included in the metadata whenever this information was available.

Besides planktonic foraminifera Mg/Ca ratios, the U\(^{37}\) index can provide information about near sea surface temperature. The U\(^{37}\) index is based on the relative degree of unsaturation of C\(_x\) alkenones, which is linearly related to temperature (Prahl et al., 1988). Alkenones are produced by coccolithophores, marine phytoplankton living in the photic zone. The production of alkenones is in many regions not constant during the year, thus potentially causing a seasonal recording bias in the U\(^{37}\) temperature proxy (Rosell-Méline and Prahl, 2013). Several calibrations exist that relate the index to sea surface temperature and if the calibration was mentioned in the original publication it was preserved in the metadata.

A large proportion of the temperature estimates in this data product are based on microfossil (planktonic foraminifera, diatoms, radiolaria, dinoflagellate cysts) assemblages. These reconstructions are based on a statistical relationship between species assemblages and temperature (Imbrie and Kipp, 1971). In theory, microfossil assemblages can be used to reconstruct temperatures of different seasons, or different environmental parameters from the same assemblage. However, it is not always clear that such reconstructions are truly independent (Telford and Birks, 2011). Several different methods exist to relate fossil assemblages to temperature and researchers often apply more than a single method in their reconstructions to increase confidence (Kucera et al., 2005). When available, these different reconstructions are included in the data product.

The bulk sediment data (CaCO\(_3\), TOC and BSI) form a category of their own. They are not proxies in the strict sense, but properties of the sediment that reflect a combination of export productivity, sedimentation and preservation. However, they can provide crucial information about the ocean-climate
system, in particular on biogeochemical cycles (Cartapanis et al., 2016). With the advent of explicit sediment modules in climate models (Heinze et al., 1999; Kurahashi-Nakamura et al., 2019), sediment composition can also be directly compared with model output and potentially provide additional constraints on the simulations.

4. Structure of the database

Following the data synthesis strategy outlined above, we generated a first data product for time series of eight parameters in sediment cores with radiocarbon and benthic δ¹⁸O stratigraphy. Following the logic of our approach, the synthesis is organised by the physical object from which the records were extracted (cores), here called site, to account for inclusion of records from spliced cores. Each site in the data product has information on seven different themes (Fig. 2):

1. **Geographic** data contains the site name, latitude and longitude (in decimals N and E) and elevation/water depth (in m) and possible notes that are relevant to the site/core as a whole. All fields except notes are essential and always included.

2. **Metadata** includes the original parameter name as given in the online data file, a standardised parameter name, parameter type (measured or inferred), unit, an estimate of analytical error as determined from repeat measurements of a standard and of reproducibility as determined by repeat measurements on samples. For measured parameters information on the instrument and laboratory is given. All metadata terms are listed in Table 2 and an overview of the standardised parameter names is provided in Table 3.

3. **Chronology** data contains raw data on absolute age control points used for age modelling. This includes depth, radiocarbon ages including their uncertainty, dated material, laboratory codes, but also calendar ages of tephra layers and palaeomagnetic events. Age control points not used in the age model by the authors of the original publication(s) are indicated and if available the source (DOI or URL) of the data is shown in addition to the original publication DOI. A complete list of chronology data terms is given in Table 4.

4. The actual time series data are provided on a common depth scale. Original age models are preserved alongside the data time series as they may differ for different time series from the same site. The data also contains information on sample number/label (mainly for DSDP/ODP/IODP cores) for spliced records and sample specific notes.

5. The **revised age model** contains ages for depths bracketed by age control points (absolute and relative). Mean and median ages are given as well as an uncertainty range (2.5 and 97.5 percentiles) based on the full suite of age model ensembles. No attempts were made to extrapolate the age models beyond the tie points in the 130,000-year time frame, so original age models may extend to either side.
6. The BACON data contains all information to reproduce the revised age model. Besides the $^{14}$C and absolute age control points, this includes the tie points for the alignment, the alignment target and all parameters used to construct the age-depth model.

7. Age ensembles are provided for further assessment of chronological uncertainty. In order to keep file sizes manageable, 1,000 randomly selected age model ensembles are preserved.

5. PALMOD 130k marine palaeoclimate data synthesis v1.0 contents

The data product contains 896 time series of the palaeoclimate parameters listed in Table 1 from 143 sites. By design all sites have both benthic stable oxygen isotope and radiocarbon data. The majority of the sites are close to the continents and in the northern hemisphere, with a concentration in the North Atlantic Ocean (Fig. 3). This reflects both research attention as well as the challenges of obtaining sediment cores with high accumulation rates and well-preserved foraminifera. The effect of research focus is also visible in the temporal coverage of the time series, where every parameter is characterised by a clear maximum around the last glacial termination ca. 15 ka BP (Fig. 4). The median resolution of the time series varies by two orders of magnitude, but is generally better than 1 sample per 1,000 years and fairly similar among the different parameters (Fig. 5). The updated age models are based on chronological control points (tie points) from radiocarbon dating, absolutely dated layers (using tephra and/or palaeomagnetic event stratigraphy) as well as alignment to the regional benthic $\delta^{18}$O stacks. The majority of the time series have a chronological control point at least every 5,000 years (Fig. 6). Taken together, the coverage in space, time and across parameters indicate that the PALMOD marine palaeoclimate data product allows for analysis of palaeoclimate on a supra-regional scale over the entire 130,000-year time frame.

This data product builds upon previous syntheses. Virtually all of the sites are also part of the benthic foraminifera $\delta^{18}$O and $\delta^{13}$C compilations of Lisiecki and Stern (2016) and Peterson and Lisiecki (2018). Our synthesis however, also includes data on other palaeoclimate parameters and contains more metadata and information on the age-depth models. Some of the planktonic foraminifera $\delta^{18}$O and Mg/Ca time series in the PALMOD 130k data product are also included in the iso2k synthesis effort (Konecky et al., 2018) and a number of sea surface temperature time series are also part of the forthcoming temperature12k synthesis (Kaufman et al., submitted).

6. Data formats and access

We provide the data products in three different formats in order to facilitate access and analysis using different software and across operating systems. Given the structure of the formats, each representation...
is slightly different in its level of metadata detail and the way metadata and data are stored. The differences are described below.

- Since the data product was built using R the data product is presented in R-readable RDS files that contain for each site a list with data for each theme (section 4). This is the format that is most complete, yet in the interest of memory space it preserves a random selection of 1,000 age models from the larger ensemble produced using BACON. In this format, all data and metadata for each site are contained within a single file. Sample scripts (https://github.com/lukasjonkers/PALMODutils) allow the user to extract a quick overview of the contents of the data product similar to Table 5, but with additional information on the temporal range, resolution and age control of the time series. Additional code is available to query the data product by parameter, parameter detail, sensor species, temporal range, resolution and age control.

- The data product is also provided in the LiPD format, which is built around JSON-LD and csv formats and is widely readable across different platforms. As for RDS, all data for each site is presented in a single file. Utilities to interact with LiPD files in R, Matlab and Python are available at https://github.com/nickmckay/LiPD-utilities.

- Finally, the data is also provided in netCDF format in a way that allows reading in PDV. This means that a single site has separate files for each individual palaeoclimate parameter as well as for the age model. In this format, some metadata is stored as concatenated strings, rather than easily searchable attributes. The netCDF format allows however to store the full suite of age model ensembles without excessive file sizes. The PDV software to read and process the data can be downloaded at https://www.marum.de/en/Stefan-Multiza/PaleoDataView.html.

The data product is available for download at https://doi.pangaea.de/10.1594/PANGAEA.908831 (Jonkers et al., 2019). We encourage users of the data product also to cite the primary source of the data when using (individual time series of) this product.

### 7. Future plans and versioning

To increase the spatio-temporal coverage over the entire 130,000-year time frame of the database, updates of this data product will first aim for quantitative growth of the database by adding more time series with chronological control based on benthic foraminifera δ¹⁸O and absolute age control points other than ¹⁴C. If available these updates will also include the parameters listed in Table 1. They will be named using the counter following the first decimal separator. The structure of the data product is designed to be flexible, allowing for the addition of different metadata fields and parameters. Further updates that include new parameters and/or require a new age modelling approach (i.e. no benthic δ¹⁸O alignment) will be named using the counter before the first decimal separator. Any updates to add or
correct (meta)data to an existing version and that do not increase the number of sites will be indicated using a counter following the second decimal separator.

8. Lesson learned: recommendations for data archiving

Data re-use and sharing are both made easier when data are archived in a standardised manner. Even though a large amount of palaeoceanographic data is publicly available, metadata to facilitate interpretation of the raw or inferred palaeodata is often not, or only partially, made available and needs to be obtained from the original publication. Synthesis efforts therefore still require a lot of time and effort to find, compile and standardise data and metadata. Only recently has the palaeoclimate community started to discuss data archiving standards (Khider et al., 2019). However, implementation of the proposed Paleoclimate Community reportIng Standard (PaCTS 1.0) will only affect new uploads to public repositories and data already available (legacy data) is likely only going to be made compliant with PaCTS through dedicated synthesis efforts. Below we list some of the main issues that we encountered during data synthesis. Our aim with mentioning these is to raise awareness of how the lack of standardisation affects data synthesis and thereby to encourage best practice in data reporting. We encourage researchers and also reviewers to treat data handling not as an afterthought, but as an integral part of their study. After all, compared to generating the data, data handling is not a time-consuming task. Time spent on proper documenting and archiving is not wasted as it facilitates reuse of data and enables scientific progress in our field.

Disambiguate core names: An apparently trivial, but surprisingly common, first-order issue is that core names are inconsistently archived. Different names for the same core arise from differences in hyphenation, truncation of (long) names, minor variations of the same name that can, with expert knowledge, be linked, but also the use of altogether different names for the same core, e.g. reflecting differences in the labelling during an expedition and in the repository. This naming confusion renders it difficult to combine data sets from the same core, especially in an automated way, and to assess the uniqueness of time series from the same core for dereplication. We recommend to use the full name as indicated in the cruise report where the core was first described.

Standardise vocabularies: Even though a vast amount of palaeoclimate data is available in public repositories, the lack of standardisation of parameter vocabularies hinders efficient data processing. This problem is clearly illustrated by the fact that for this synthesis long synonym lists needed to be generated in order to group parameters. Each parameter in this synthesis had 10s-100s of different names, of which many were unique (Fig. 7). This issue can be partly addressed by a consistent separation of parameter and attribute names (e.g. parameter: δ¹⁸O, species: G. bulloides, instead of a single parameter “d18OGBu”), but even that calls for a standardisation of parameters and attributes.
Report sampling depth: All data reported here are based on measurements on discrete samples from a specific depth interval in a given core. Therefore, the synthesis requires information on the position of each sample. Since ages of the samples are always estimates and may differ among studies of the same archive, unambiguous information on sample depth is essential to reproduce and update the time series. Despite this, many studies fail to report sample depth and instead report only age. This problem is worse for spliced records that rely on a composite depth scale. Splicing approaches are often opaque and original sample depths, or sample codes that identify unique samples, are not always available. We recommend therefore that sample depth should be essential for palaeodata time series from (marine) sediments and that sample labels are archived for spliced records. This includes (I)OPD/DSDP sample labels (in full) or IGSN (if available).

Publish raw data: To ensure reproducibility of inferred parameters (in this synthesis only temperature) and to ensure harmonisation or updating of the calibration, the raw measured data are needed. Provided that the calibration is known, measured data can in some cases be calculated from the inferred data, but this is impossible for data based on microfossil transfer functions. Related to this is unclear information about the calibration that was used, particularly if the publication describing the calibration includes multiple different equations.

Include metadata: To assess ecological imprints on proxy signals (recording bias), temperature estimates as well as oxygen and carbon isotope data based on planktonic foraminifera also require that key metadata, such as species name are archived in a standardised way. This is not universally done and for example the species information is often only available in the original publication. Additional information to assess the uncertainty of proxy measurements, such as foraminifera shell size, the sample size (e.g. number of shells, concentration of alkenones), reproducibility of repeat measurements was often not available from the paper nor from the archived data and we encourage the archiving of such data in a standardised way. A similar issue applies to chronological data. Thanks to a longer history of reporting standards, radiocarbon (Stuiver and Polach, 1977) (meta)data is often rather complete. However, this information is often not included alongside the digitally available data and for this synthesis a large part of the radiocarbon (meta)data had to be scraped from literature. In our age modelling approach, we took reservoir age uncertainty into account, using data derived from the modelled reservoir ages (see section 2.3). Alternative approaches are hindered by the fact that reservoir age uncertainty is almost never reported.

Avoid redundancy: A considerable amount of time was spent on dereplication of time series of the same parameter from the same core that were archived multiple times. Repeat archiving happens when data
are reused or, less commonly, updated. The dereplication task is not made easier by (incomplete or inconsistent) metadata reporting and can be avoided through better linking of existing data sets, when instead of re-uploading the data the DOI/URL of the original data is provided.

Help rescue dark data: This synthesis is based on data that are publicly available, yet many/some palaeoceanographic time series are not archived in public repositories. Even though the proportion of this so-called dark data is unknown, they likely affect every branch of palaeoceanography and as a result palaeoceanographical data syntheses cannot be exhaustive. This problem is clearly exacerbated for syntheses relying on automated data mining. There are several shades of dark data, each requiring their own approach to retrieve and make them available. Some data is only partially available, for instance data sets that lack sample depths. Such data only require additional data to make them re-useable. This additional data can sometimes be calculated, or obtained from cross-referencing different data files, but in many cases will need to be retrieved from the data producers. Other data sets, in particular those from before the digital age, are presented in tables in the original publications. Progress has been made with digitising those data sets (especially PANGAEA), but this work is not finished and more effort is needed to make this data available to the community. There is also data that is used in publications, but is not made available in any way (print or digital). This third shade of data can so far only be obtained from the original data producers or authors of the original publication, or if this proves impossible, needs to be digitised from graphs. Digitisation inevitably leads to loss of accuracy of the data and a data set retrieved in this way should be flagged. A final category of dark data consists of data that is not part of a publication. Such data sets can be made publicly available and be associated with a DOI to ensure traceability. To reward data sharing the use of data citations needs to be encouraged and data citations should be included in the evaluation of a researcher’s impact.

Acknowledgments

This research was supported by PALMOD, the German palaeoclimate modelling initiative (www.palmod.de). PALMOD is part of the Research for Sustainability initiative (FONA; www.fona.de) funded by the German Federal Ministry of Education and Research (BMBF). We thank colleagues in working group 3 from PALMOD for discussion on data and metadata requirements and Rangelys Sorrentino for help with data processing. We also thank Blanca Ausin and an anonymous reviewer for their constructive feedback on an earlier version of this manuscript.

Data availability statement

An overview of all datasets used in this synthesis, including URLs to the data can be found in Table 5 (https://seafile.zfn.uni-bremen.de/f/a5732554bc91413780a3/). The PALMOD 130k marine palaeoclimate
The data product can be downloaded in R, LiPD and netCDF format at https://doi.pangaea.de/10.1594/PANGAEA.908831. The data can also be visualised and downloaded in LiPD and csv formats at http://lipdverse.org/PalMod/current_version/.
### Tables

Table 1: Palaeoclimate parameters in the PALMOD 130k marine data synthesis.

<table>
<thead>
<tr>
<th>Parameter</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Benthic foraminifera δ¹⁸O and δ¹³C</td>
<td></td>
</tr>
<tr>
<td><strong>Planktonic</strong> foraminifera δ¹³C</td>
<td></td>
</tr>
<tr>
<td>Seawater temperature*</td>
<td></td>
</tr>
<tr>
<td>Radiocarbon</td>
<td></td>
</tr>
<tr>
<td>Carbonate content</td>
<td></td>
</tr>
<tr>
<td>Total organic carbon content</td>
<td></td>
</tr>
<tr>
<td>Biogenic silica content</td>
<td></td>
</tr>
</tbody>
</table>

* inferred from various proxies (foraminifera Mg/Ca, alkenones, microfossil assemblages)
<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ParameterOriginal*</td>
<td>Original parameter name as in data repository</td>
</tr>
<tr>
<td>Parameter*</td>
<td>Standardised parameter name (see Table 4)</td>
</tr>
<tr>
<td>ParameterType*</td>
<td>Parameter type, measured or inferred</td>
</tr>
<tr>
<td>ParameterUnit*</td>
<td></td>
</tr>
<tr>
<td>ParameterAnalyticalError</td>
<td>Error based on repeat measurements of standards</td>
</tr>
<tr>
<td>ParameterReproducibility</td>
<td>Error based on repeat measurements of samples</td>
</tr>
<tr>
<td>Instrument</td>
<td></td>
</tr>
<tr>
<td>Laboratory</td>
<td></td>
</tr>
<tr>
<td>SampleThickness_cm</td>
<td></td>
</tr>
<tr>
<td>Material*</td>
<td>Measurement material or parameter on which inferred parameter is based</td>
</tr>
<tr>
<td>Species</td>
<td></td>
</tr>
<tr>
<td>Nshells</td>
<td></td>
</tr>
<tr>
<td>SizeFraction_microm</td>
<td></td>
</tr>
<tr>
<td>Notes</td>
<td></td>
</tr>
<tr>
<td>RecordingSeason</td>
<td></td>
</tr>
<tr>
<td>RecordingDepth</td>
<td></td>
</tr>
<tr>
<td>EquilibriumOffset</td>
<td></td>
</tr>
<tr>
<td>CalibrationEquation</td>
<td></td>
</tr>
<tr>
<td>CalibrationUncertainty</td>
<td></td>
</tr>
<tr>
<td>CalibrationDOI</td>
<td></td>
</tr>
<tr>
<td>TransferFunctionTrainingSet</td>
<td></td>
</tr>
<tr>
<td>TransferFunctionUncertainty</td>
<td></td>
</tr>
<tr>
<td>TransferFunctionDOI</td>
<td></td>
</tr>
<tr>
<td>PublicationDOI*</td>
<td></td>
</tr>
<tr>
<td>Authors</td>
<td></td>
</tr>
<tr>
<td>PublicationTitle</td>
<td></td>
</tr>
<tr>
<td>Journal</td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td></td>
</tr>
<tr>
<td>Volume</td>
<td></td>
</tr>
<tr>
<td>Issue</td>
<td></td>
</tr>
<tr>
<td>Pages</td>
<td></td>
</tr>
<tr>
<td>ReportNumber</td>
<td>DataDOI</td>
</tr>
<tr>
<td>--------------</td>
<td>---------</td>
</tr>
<tr>
<td>Name</td>
<td>Description</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>benthic.d18O</td>
<td>Benthic foraminifera δ¹⁸O</td>
</tr>
<tr>
<td>benthic.d13C</td>
<td>Benthic foraminifera δ¹³C</td>
</tr>
<tr>
<td>planktonic.d18O</td>
<td>Planktonic foraminifera δ¹⁸O</td>
</tr>
<tr>
<td>planktonic.d13C</td>
<td>Planktonic foraminifera δ¹³C</td>
</tr>
<tr>
<td>surface.temp</td>
<td>Inferred (near) sea surface temperature (based on microfossils, planktonic foraminifera Mg/Ca, U³⁷)</td>
</tr>
<tr>
<td>deep.temp</td>
<td>Inferred bottom water temperature (based on benthic foraminifera Mg/Ca)</td>
</tr>
<tr>
<td>CaCO3</td>
<td>Calcium carbonate content</td>
</tr>
<tr>
<td>TOC</td>
<td>Total organic carbon content</td>
</tr>
<tr>
<td>BSi</td>
<td>Biogenic silica content</td>
</tr>
<tr>
<td>DBD</td>
<td>Dry bulk density</td>
</tr>
<tr>
<td>IRD</td>
<td>Ice-rafted detritus</td>
</tr>
<tr>
<td>planktonic.MgCa</td>
<td>Planktonic foraminifera Mg/Ca ratio</td>
</tr>
<tr>
<td>benthic.MgCa</td>
<td>Benthic foraminifera Mg/Ca ratio</td>
</tr>
<tr>
<td>UK37</td>
<td>U³⁷ ratio (rare cases where this is not U³⁷ mentioned in notes)</td>
</tr>
<tr>
<td>C37.concentration</td>
<td>Alkenone concentration</td>
</tr>
</tbody>
</table>
Table 4: Chronology terms, asterisks mark essential terms.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ChronType*</td>
<td>Type of absolute chronology tie point ((^{14})C, tephra, palaeomag)</td>
</tr>
<tr>
<td>ChronDepthTop_cm</td>
<td></td>
</tr>
<tr>
<td>ChronDepthBottom_cm</td>
<td></td>
</tr>
<tr>
<td>ChronDepthMid_cm*</td>
<td></td>
</tr>
<tr>
<td>ChronSampleThickness_cm</td>
<td></td>
</tr>
<tr>
<td>ChronAge_kaBP*</td>
<td>Age of non-(^{14})C tie point (tephra, palaeomag)</td>
</tr>
<tr>
<td>ChronAgeError_ka*</td>
<td>Age error of non-(^{14})C tie point (tephra, palaeomag)</td>
</tr>
<tr>
<td>ChronDatedMaterial</td>
<td></td>
</tr>
<tr>
<td>ChronDatedSpecies</td>
<td></td>
</tr>
<tr>
<td>ChronNshellsDated</td>
<td></td>
</tr>
<tr>
<td>Chron14CLabcode</td>
<td></td>
</tr>
<tr>
<td>ChronAge14C_kaBP*</td>
<td></td>
</tr>
<tr>
<td>ChronAge14CError_ka*</td>
<td></td>
</tr>
<tr>
<td>ChronAge14CErrorUp_ka</td>
<td></td>
</tr>
<tr>
<td>ChronAge14CErrorDown_ka</td>
<td></td>
</tr>
<tr>
<td>ChronReservoirAge_ka*</td>
<td></td>
</tr>
<tr>
<td>ChronReservoirAgeError_ka*</td>
<td></td>
</tr>
<tr>
<td>ChronCalibCurve</td>
<td></td>
</tr>
<tr>
<td>ChronCalibAge14C_kaBP</td>
<td>Calibrated (^{14})C age</td>
</tr>
<tr>
<td>ChronCalibAge14CI1sigLo_ka</td>
<td></td>
</tr>
<tr>
<td>ChronCalibAge14CI1sigUp_ka</td>
<td></td>
</tr>
<tr>
<td>ChronAgeModelMethod</td>
<td></td>
</tr>
<tr>
<td>ChronAgeRejected</td>
<td></td>
</tr>
<tr>
<td>ChronNotes</td>
<td></td>
</tr>
<tr>
<td>ChronSource</td>
<td></td>
</tr>
<tr>
<td>ChronDOI*</td>
<td></td>
</tr>
</tbody>
</table>
Table 5 (online only https://seafile.zfn.uni-bremen.de/f/a5732554bc91413780a3/; will be made available on zenodo after peer review): Palaeoclimate time series in the PALMOD 130k marine palaeoclimate data synthesis v1.0.1. This table lists the site names and location, parameters including additional information as well as the source of the data and the original publications where the data were presented.
Figure 1: Data synthesis approaches. In the expansion approach the database size increases slowly as records are added. Database size follows an opposite pathway using reduction approach and reaches a stable size quicker, with less effort. Since the expansion approach is not restricted to data that is available in the public domain, this approach may lead to a database that includes data that is not publicly available (dark data). The reduction approach on the other hand is arguably more objective, can be automated and is therefore more efficient. This approach also encourages good data stewardship.
Figure 2: Structure of the PALMOD 130k marine palaeoclimate data synthesis. Each file in the database contains information on seven different themes on a single site (sediment core). Links between different themes are indicated.
Figure 3: Spatial distribution of sites in version 1.0 of the PALMOD 130k marine palaeoclimate data synthesis. The distribution of the sites reflects research effort and the possibility to obtain sediment cores containing well-preserved foraminifera and is hence skewed towards the northern hemisphere (Atlantic) and the continental margins. Since no explicit search was done for other parameters than benthic foraminifera δ¹⁸O, the distribution of the other parameters is restricted to sites that have benthic foraminifera δ¹³C. For benthic foraminifera δ¹³C only sites with data based on the genus Cibicidoides are shown.
Figure 4: Temporal distribution of time series in version 1.0 of the PALMOD 130k marine palaeoclimate data synthesis. The number of time series (y axis) is counted per 1,000 year bin. The temporal availability of all parameters shows a clear maximum around 15 ka BP, reflecting the research focus on the last glacial termination. For benthic foraminifera δ\textsuperscript{13}C only time series with data based on the genus Cibicidoides are shown.
Figure 5: median resolution of the time series in version 1.0 of the PALMOD 130k marine palaeoclimate data synthesis. Box-whisker plots show spread of resolution per parameter. For benthic foraminifera δ¹³C the resolution data are restricted to time series containing data measured on shells of the genus Cibicidoides. For all parameters the median resolution is more than 1 data point per 1,000 years.
Figure 6: Average temporal spacing between chronology tie points (AvgPoints). Tie points are also split in $^{14}$C, absolute (tephra/palaeomagnetics) and tuned. The majority of the time series in the synthesis have a spacing of chronology tie points that is below 5,000 years.
Figure 7: The problem of absence of standardisation in parameter names. The cumulative frequency of synonyms for seawater temperature in our initial data base (section 2.2.1), showing that there are over 500 different names for the same parameter and that many of these are unique.
References


Broecker, W. S., Andre, M., Bonani, G., Wolff, W., Klas, M., Mix, A., and Oeschger, H.: Comparison between radiocarbon ages obtained on coexisting planktonic foraminifera, Paleoceanography, 3, 647-657, 10.1029/pa003i006p00647, 1988a.


western Iberian margin during key cold episodes of the last 45 ka, Geochemistry, Geophysics, Geosystems, 10, n/a-n/a, 10.1029/2009gc002398, 2009.


Kienast, S. S., Calvert, S. E., and Pedersen, T. F.: Nitrogen isotope and productivity variations along the northeast Pacific margin over the last 120 kyr: Surface and subsurface paleoceanography, Paleoceanography, 17, 7-1-7-17, 10.1029/2001pa000650, 2002.


Latif, M., Claussen, M., Schulz, M., and Brüchner, T.: Comprehensive Earth system models of the last glacial cycle, EOS, 97, 10.1029/2016EO059587, 2016.


Marcott, S. A., Shakun, J. D., Clark, P. U., and Mix, A. C.: A Reconstruction of Regional and Global Temperature for the Past 11,300 Years, Science, 339, 1198-1201, 10.1126/science.1228026, 2013.


Nam, S.-I.: Late Quaternary glacial history and paleoceanographic reconstructions along the East Greenland continental margin: Evidence from high-resolution records of stable isotopes and ice-rafted debris (Spätquartäre Vereisungsgeschichte und paläozeanographische Rekonstruktionen am ostgrönlandischen Kontinentalrand), Alfred-Wegener-Institut für Polar- und Meeresforschung0176-5027, 1997.


Oppo, D. W., and Sun, Y.: Amplitude and timing of sea-surface temperature change in the northern South China Sea: Dynamic link to the East Asian monsoon, Geology, 33, 785, 10.1130/g21867.1, 2005.


Pichevin, L., Martinez, P., Bertrand, P., Schneider, R., Girardeau, J., and Eneas, K.: Nitrogen cycling on the Namibian shelf and slope over the last two climatic cycles: Local and global forcings, Paleoceanography, 20, n/a-n/a, 10.1029/2004pa001001, 2005.


Praetorius, S. K., McManus, J. F., Oppo, D. W., and Curry, W. B.: Episodic reductions in bottom-water currents since the last ice age, Nature Geoscience, 1, 449-452, 10.1038/ngeo227, 2008.


Rickaby, R. E. M., and Elderfield, H.: Evidence from the high-latitude North Atlantic for variations in Antarctic Intermediate water flow during the last deglaciation, Geochemistry, Geophysics, Geosystems, 6, n/a-n/a, 10.1029/2004gc000858, 2005.


Saikku, R., Stott, L., and Thunell, R.: A bi-polar signal recorded in the western tropical Pacific: Northern and Southern Hemisphere climate records from the Pacific warm pool during the last Ice Age, Quaternary Science Reviews, 28, 2374-2385, 10.1016/j.quascirev.2009.05.007, 2009.


Sarnthein, M., Schneider, B., and Grootes, P. M.: Peak glacial ¹⁴C ventilation ages suggest major draw-down of carbon into the abyssal ocean, Climate of the Past, 9, 2595-2614, 10.5194/cp-9-2595-2013, 2013.


Shakun, J. D., Clark, P. U., He, F., Marcott, S. A., Mix, A. C., Liu, Z., Otto-Bliesner, B., Schmittner, A., and Bard, E.: Global warming preceded by increasing carbon dioxide concentrations during the last deglaciation, Nature, 484, 49-54, 2012.


Snyder, C. W.: Evolution of global temperature over the past two million years, Nature, 538, 226-228, 10.1038/nature19798, 2016.


Vidal, L., Schneider, R. R., Marchal, O., Bickert, T., Stocker, T. F., and Wefer, G.: Link between the North and South Atlantic during the Heinrich events of the last glacial period, Climate Dynamics, 15, 909-919, 10.1007/s003820050321, 1999.


Voelker, A. H. L., de Abreu, L., Schönfeld, J., Erlenkeuser, H., and Abrantes, F.: Hydrographic conditions along the western Iberian margin during marine isotope stage 2, Geochemistry, Geophysics, Geosystems, 10, n/a-n/a, 10.1029/2009gc002605, 2009.


Wells, P., and Okada, H.: Response of nannoplankton to major changes in sea-surface temperature and movements of hydrological fronts over Site DSDP 594 (south Chatham Rise, southeastern New Zealand), during the last 130 kyr, Marine Micropaleontology, 32, 341-363, 10.1016/s0377-8398(97)00025-x, 1997.


