https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021


Open Access Earth System Science Data Discussions

# Description of a global marine particulate organic carbon-13 isotope data set

Maria-Theresia Verwega[1, 2], Christopher J. Somes[1], Markus Schartau[1], Robyn E. Tuerena[3], Anne Lorrain[4], Andreas Oschlies[1], and Thomas Slawig[2]

[1]GEOMAR - Helmholtz Centre for Ocean Research, Kiel, Germany
[2]Kiel University, Kiel, Germany
[3]Scottish Association for Marine Science, Dunstaffnage, Oban PA37 1QA
[4]Univ Brest, CNRS, IRD, Ifremer, LEMAR, F-29280 Plouzané, France

**Correspondence:** mverwega@geomar.de

**Abstract.** Marine particulate organic carbon-13 stable isotope ratios ($\delta^{13}C_{POC}$) provide insights in understanding carbon cycling through the atmosphere, ocean, and biosphere. They have been used to trace the input of anthropogenic carbon in the marine ecosystem due to the distinct isotopically light signature of anthropogenic emissions. However, $\delta^{13}C_{POC}$ is also significantly altered during photosynthesis by phytoplankton, which complicates its interpretation. For such purposes, robust spatio-temporal coverage of $\delta^{13}C_{POC}$ observations is essential. We collected all such available data sets, merged and homogenized them to provide the largest available marine $\delta^{13}C_{POC}$ data set (Verwega et al., 2021). The data set consists of 4732 data points covering all major ocean basins beginning in the 1960s. We describe the compiled raw data, compare different observational methods, and provide key insights in the temporal and spatial distribution that is consistent with previously observed patterns. The main different sample collection methods (bottle, intake, net, trap) are generally consistent with each other when comparing within regions. An analysis of 1990s mean $\delta^{13}C_{POC}$ values in an meridional section accross the Atlantic Ocean shows relatively high values ($\geq -22‰$) in the low latitudes ($< 30°$) trending towards lower values in the Arctic Ocean ($\sim -24‰$) and Southern Ocean ($\leq -28‰$). The temporal trend since the 1960s shows a decrease of mean $\delta^{13}C_{POC}$ by more than $3‰$ in all basins except for the Southern Ocean which shows a weaker trend but contains relatively poor multi-decadal coverage.

## 1 Introduction

Carbon is an essential element for life and it is regulating climate via atmospheric form $CO_2$, a long-living greenhouse gas. Understanding carbon cycling is fundamental to reliably project changes of the Earth's future climate. Carbon is subject to transformation and cycling throughout the ocean, land and atmosphere. It is a major part of organic matter of all living organisms which can both consume (e.g. photosynthesis) and produce (e.g. respiration) inorganic carbon. Besides the natural cycling processes, the total amount and distribution of carbon is strongly perturbed by human activity caused by the industrialization, most notably due to fossil fuel emissions, deforestation, farming, cement production and other industrial processes. Anthropogenic $CO_2$ emissions are one of the main driving forces of modern climate change which is likely to continue in the

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021


Open Access Earth System Science Data Discussions

future (IPCC, 2013). Only about 60 % of the anthropogenic $CO_2$ emissions have been compensated by natural sinks, including the dissolution of inorganic carbon in the ocean. This leaves the atmosphere enriched with anthropogenic carbon already by about 880 Gt $CO_2$ since 1750 (IPCC, 2014), which is driving the increase of global temperature levels. The ocean serves as an important buffer, as it absorbs a significant amount of anthropogenic carbon, with the ocean interior being the largest readily exchangeable reservoir of carbon in the Earth system.

Marine phytoplankton convert dissolved inorganic carbon (e.g. aqueous $CO_2$) into their organic carbon via photosynthesis in the euphotic surface layer. This organic carbon forms the base of the food web for higher tropic levels in marine ecosystems. Parts of the particulate organic carbon (POC) sinks down to ocean depths, where it is either respired back to dissolved inorganic carbon by heterotrophic organisms or becomes buried in ocean sediments (Suess, 1980). This process is known as the soft-tissue biological carbon pump, an important mechanism for sequestering carbon to the deep ocean from the atmosphere (Rocha and Passow, 2014). Since the deep ocean has a residence time of about a millennium, it is a key carbon reservoir influencing long-term climate change.

Carbon isotopes provide additional insights into the cycling of carbon in the Earth system. The element carbon exists in two naturally occurring stable isotopes, $^{12}$C and $^{13}$C, with abundances of around 98.9 % and 1.1 %, respectively. Knowledge of their pathways through carbon reservoirs can support deeper understanding of carbon transfer and can help identify carbon sources with different isotopic ratios (Rounick and Winterbourn, 1986). Relative abundances of carbon isotopes are usually given as the $\delta$-notation, which is based on the carbon isotope ratio $\frac{^{13}\mathrm{C}}{^{12}\mathrm{C}}$, standardized and given in parts per thousands as

$$\delta^{13}\mathrm{C} = \left( \frac{\frac{^{13}\mathrm{C}}{^{12}\mathrm{C}}}{R_{std}} - 1 \right) 1000. \tag{1}$$

The constant $R_{std} = 0.0112372$ is a standard ratio, originally referring to the calcareous fossil PeeDee Belmnite. The values $^{12}$C and $^{13}$C are the absolute concentrations of the individual isotopes (Hayes, 2004).

Phytoplankton preferentially incorporate (i.e. fractionate) the lighter $^{12}$C carbon isotope into its organic matter. This fractionation causes phytoplankton organic $\delta^{13}$C to be 10 to 25 ‰ lower than that of inorganic $\delta^{13}$C, which depends on a variety of environmental, ecological, and physiological conditions (e.g. Popp et al., 1989, 1998; Rau et al., 1989, 1996). The main factors that control phytoplankton fractionation are concentrations of $CO_2\,[aq]$, species-specific effects enforced by the phytoplankton composition, and cellular growth rate, although their are large uncertainties regarding the specific processes and mechanisms that cause variations in phytoplankton fractionation (e.g. Fry, 1996; Laws et al., 1995; Popp et al., 1998; Bidigare et al., 1997; Cassar et al., 2006).

$\delta^{13}C_{POC}$ provides insights into physical and biological carbon cycle processes in the ocean (e.g. Fry and Sherr, 1989). It helps to diagnose carbon pathways from the atmosphere to the deep ocean including the biological carbon pump (e.g. Jasper and Hayes, 1990; Popp et al., 1989; Freeman and Hayes, 1992), assists reconstruction of oceanic carbon cycling and even plankton cell sizes and communities (e.g. Tuerena et al., 2019; Lorrain et al., 2020). For example, anthropogenic carbon emissions have a distinctly low $\delta^{13}$C content, making $\delta^{13}$C a useful property for tracing anthropogenic carbon throughout the Earth system. Atmospheric $\delta^{13}C_{CO_2}$ has decreased from -6.5 ‰ in preindustrial times to -8.6 ‰ presently. The measurable decrease due to anthropogenic fossil carbon emissions is known as the Suess Effect (Keeling, 1979), which enters the ocean via air-sea

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021




gas exchange. However, since changes in marine $\delta^{13}C_{POC}$ are also significantly influenced by changes in phytoplankton fractionation due to other anthropogenic controls (e.g. increasing $CO_2\,[aq]$ concentrations, changes in temperature-dependent growth rates, changing phytoplankton communities), determining the driving processes(es) of $\delta^{13}C_{POC}$ spatial and temporal trends remains a challenge. A better understanding of the contributions from all of these effects requires a robust global data set of $\delta^{13}C_{POC}$.

Models can be used to describe, project, and understand changes associated with $\delta^{13}C_{POC}$. Earth system models serve to simulate and test hypotheses in different scenarios as unbiased assessments (e.g. IPCC, 2014) and may support future decision making. Besides resolving mass flux of carbon, many models also simulate stable carbon isotopes (e.g. Schmittner and Somes, 2016; Buchanan et al., 2019; Hofmann et al., 2000; Jahn et al., 2015; Tagliabue and Bopp, 2008; Morée et al., 2018). For reliable calibrations of models, a spatially and temporally comprehensive data set is essential. This additional constraint provided by marine $\delta^{13}C_{POC}$ assists reconstruction of oceanic carbon cycling including how much anthropogenic carbon is entering marine ecosysems and exported to the deep ocean.

Data sets of marine $\delta^{13}C_{POC}$ improve our understanding marine carbon cycling by providing another independent constraint. To date, numerous individual $\delta^{13}C_{POC}$ data sets exist, while the number of accessible, merged data sets is lacking. Existing merged data sets contain data from several sources but were often focused on a specific region or process. Individual data sets are usually collected during a specific cruise or time series station and are often neglected since they contain relatively few data. Such data sets can easily be accessed on data platforms such as PANGAEA and, when combined, they can represent an important and significant source of data.

In this study, we provide a novel merged seawater $\delta^{13}C_{POC}$ data product (Verwega et al., 2021), that – to our knowledge – contains the most expansive spatio-temporal coverage to date. It contains all available $\delta^{13}C_{POC}$ seawater data from PANGAEA and the merged data sets by Goericke (1994), Tuerena et al. (2019) and Young et al. (2013), as well as unpublished data from different cruises by Lorrain. No data were excluded, even if sampled at extreme locations (e.g, trenches, hydrothermal vents). Sampled metadata include sample location, time, depth and method as well as the original source. The data set is multilateral, which facilitates its applicability, e.g. the backtracking to individual data points. Provided data files are NetCDF files interpolated onto two different global grids and a csv file that includes the data anomalies with respect to their overall mean together with all corresponding available meta information.

The paper is structured as follows: We provide a brief overview of $\delta^{13}C_{POC}$ data acquisition in section 2 and its compilation and metadata in section 3. The characteristics of the collected $\delta^{13}C_{POC}$ data are shown in section 4. We present their spatial distribution in section 5 and temporal distribution in section 6. Lastly, we provide a short summary and concluding remarks.

## 2 Data acquisition

The data set includes 4732 entries for $\delta^{13}C_{POC}$ from 185 different sources and ranges from the 1960s to the 2010s. In addition to many data sets from the data platform PANGAEA, we included unpublished data sets provided by the coauthors Tuerena

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021


Open Access Earth System Science Data Discussions

and Lorrain and the data products from Goericke (1994) and Young et al. (2013). The conducted adjustments described in the following.

### 2.1 Data sources

As a basis of our data set, we chose the 1990s data collection by Goericke (1994). This was established to investigate variations in $\delta^{13}C_{POC}$ with temperature and latitude. The $\delta^{13}C_{POC}$ sample data and measurements were conducted by investigating zooplankton, net-plankton or particulate organic matter. We cross-checked and extended this data set by looking up all available primary sources. Goericke originally included 476 of $\delta^{13}C_{POC}$ data points from 17 contributions. Largest contributions came from Fischer (1989) with 107 entries, Fontugne et al. (1991) with 97 and Fontugne and Duplessy (1981, 1978) with 78. Large extensions were possible e.g. in the Fischer (1989) and Eadie and Jeffrey (1973) data sets, incorporating more than 70 additional data points from these primary sources. With this extension, we could increase the data set to 626 data points for $\delta^{13}C_{POC}$.

We collected most data from the PANGAEA data platform, an open access online library archiving and providing geo-referenced Earth system data, hosted and monitored by the Alfred-Wegener-Institut - Helmholtz Center for Polar and Marine Research (AWI) and the Center for Marine Environmental Sciences, University of Bremen (MARUM). With the data made available therein, we could further extend the data set by additional $\approx$ 3,500 measurements of $\delta^{13}C_{POC}$. Most $\delta^{13}C_{POC}$ data from PANGAEA are associated with samples collected during the Joint Global Ocean Flux Study (JGOFS), with more than 2000 of $\delta^{13}C_{POC}$ data points. Additional 529 samples were contributions by the Antarctic Environments Southern Ocean Process Study (AESOPS), 342 by the Archive of Ocean Data (EurOBIS Data Management Team) and 279 by the SFB313 (Thiede et al., 1988).

Other collected data were provided by Tuerena and Lorrain. Tuerena provided a data contribution coming from the data set mentioned in Tuerena et al. (2019), to which we will refer as the Tuerena data set. This contains 595 data points including 501 from Young et al. (2013) and covers samples within the euphotic zone and an observation timeframe of 1964 - 2012 and overall referred to as the Tuerena data set. Moreover, we included 69 unpublished data points provided by Lorrain, covering the years 2012 - 2015 and sampled during the cruises CASSIOPEE, PANDORA, OUTPACE, NECTALIS 3 and 4, KH13. We refer to this data set as the Lorrain data set.

### 2.2 Adjustments made

All data were taken with as many details as possible from the sources and have eventually been reshaped to fit the structure described in Table 3. No rounding or cut off of detailed data was made. Averaging was applied for data points that only provided depth intervals, but not for timeframes. Longitude values were made fitting to the format $[-180°, 180°]$. Wherever possible the data was taken from its original publication. Changes made to the data by Goericke are described in Table 1, changes to all other data in Table 2.

Most data listed in the Goericke data set could be gathered from the original publications directly. Some data are not accessible from an original source, including those data labeled as "Harrison", "Hobson" and "Schell", which were included as unpublished data by personal communication in Goericke (1994). Also, we could not identify the original data sources of

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021


Open Access Earth System Science Data Discussions

"Voss (1991)" and "Sacket et al. (1966)". Data from these sources are used as provided by Goericke. All other data could be directly compared with and linked to their origin. We complemented the data with information about month, year, depth, sample method, cruise information, if applicable trap duration and a references according to Table 3. Special notes given in Goericke (1994) are conserved in our "project/cruise" named meta information. Suspicious or rounded values were adjusted to their source values.

Averaging was only applied for depth ranges, these were included as their arithmetic mean. Sample timeframes were only included when lying completely within one month and year. Sample depth given as "surface" was denoted as $1\ \mathrm{m}$.

Wherever multiple types of $\delta^{13}C_{POC}$ e.g. similar measurements based on different methods, were given within one source, we chose only one type. In Westerhausen and Sarnthein (2003), we chose the "mass spectrometer" data set because this was the originally measured one. In Trull and Armand (2013a) and in Trull and Armand (2013b), we used the "blanc corrections" data set of $\delta^{13}C$, since this set of $\delta^{13}C_{org}$ values is recommended to be considered (Trull and Armand, 2001).

The primary source of the Tuerena and Lorrain data is mentioned in our data set in the "Project/cruise" column. In the data set from Tuerena et al. (2019), this was originally labeled as "source", in the Lorrain data set as "campaign". In both data sets the Longitude was converted to $[-180°, 180°]$ from a $[0°, 360°]$ format. We used the transformation

$$Long_{new} = \begin{cases} Long_{old} - 360° & \text{for all } Long_{old} \in (180°, 360°] \\ Long_{old} & else \end{cases} \tag{2}$$

In the data of MacKenzie et al. (2019) we deleted a typo where the depth value was set equal to the negative Longitude value. We disregarded trap duration given in Voss and von Bodungen (2003), which was given as the negative value $-1$.

## 3 Content and structure of the data set

The data collection is made available in files of raw and interpolated values respectively. (Verwega et al., 2021). The raw data is a csv file that includes the anomalies of the $\delta^{13}C_{POC}$ with respect to their mean and all available meta information. The interpolated data is provided as NetCDF files on two different global grids: a $1.8° \times 3.6°$-resolution grid from a model that simulates $\delta^{13}C_{POC}$ and the $1° \times 1°$-resolution grid of the World Ocean Atlas.

### 3.1 Raw data file

The csv-format data file includes $\delta^{13}C_{POC}$ anomalies and meta information in its columns. A full description of the content, value range and coverage of the individual columns is given in Table 3. Anomalies of $\delta^{13}C_{POC}$ were calculated, based on the arithmetic mean of the full data collection. The mean was calculated and used as

$$mean_{\delta^{13}C_{POC}} = -23.955615278114315‰ \tag{3}$$

Anomalies contain all relevant information with respect to variability of the $\delta^{13}C_{POC}$ data in space and time. This way it becomes easier to analyze bias information separately, e.g. during first steps of model calibration.

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021


Open Access Earth System Science Data Discussions

**Table 1.** Changes that were introduced to data taken from Goericke (1994): The first column names the publication or author of the primary data set. The second column lists, in which part of the data we applied changes. The third and fourth column show, from what values to which they have been changed and the last columns gives the reason for this.

| data set | changed | from | to | reason |
|---|---|---|---|---|
| Degens et al. (1968) | Longitude | Goericke | source value | E, W interchanged |
| Eadie and Jeffrey (1973) | | | | |
| Fischer (1989) | | | | |
| Fontugne and Duplessy (1978) | | | | |
| Fontugne and Duplessy (1981), MD13 Osiris III | | | | |
| Francois et al. (1993) | | | | |
| Harrison[1] | | | | |
| Sacket et al. (1965) | | | | |
| Saupe et al. (1989) | | | | |
| Wada et al. (1987) | | | | |
| Eadie and Jeffrey (1973) | latitude, longitude | Goericke | source value | rounded in Goericke |
| Fischer (1989) all, but INDOMED leg-12 | | | | |
| Fontugne and Duplessy (1978) | | | | |
| Fontugne and Duplessy (1981) | | | | |
| Francois et al. (1993) | | | | |
| Sacket et al. (1965) | | | | |
| Eadie and Jeffrey (1973) | $\delta^{13}C_{POC}$ | not included | added | not included in Goericke |
| Fischer (1989) | | | | |
| Sacket et al. (1965) | | | | |
| Wada et al. (1987) | | | | |
| Fischer (1989) | $\delta^{13}C_{POC}$ | Goericke | source value | rounded in Goericke |
| Fontugne and Duplessy (1978) | | | | |
| Fontugne and Duplessy (1981) | | | | |
| Fischer (1989) | temperature | Goericke | source value | rounded in Goericke |
| Fontugne and Duplessy (1981) | | | | |
| Francois et al. (1993) | | | | |
| Sacket et al. (1965) | | | | |
| Fischer (1989) | $\delta^{13}C_{POC}$ | Goericke | deleted | not found in source |
| Fontugne and Duplessy (1978) | temperature | Goericke | deleted | not found in source |

[1]The original source was not available, but we highly suspected an error in the coordinates interchanged East and West.

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021


Open Access Earth System Science Data Discussions

**Table 2.** Changes made in other data: This table's structure is equivalent to Table 1. It refers to all changes made in general and any other than the Goericke (1994) data.

| data set | changed | from | to | reason |
|---|---|---|---|---|
| any | depth | "surface" | 1 | comparability |
| any | depth | depth range | average[1] | comparability |
| Trull and Armand (2013a) | $\delta^{13}C_{POC}$ | three available | "blank correction" | mentioned in Trull and Armand (2001) |
| Trull and Armand (2013b) | | | | |
| any using sediment traps | month, year | range | explicit value[2] | comparability |
| Chang et al. (2013) | month, year | | explicit number | just one date for trap sampling given |
| Lorrain | Project/cruise | | "campaign" | provided by Lorrain |
| Tuerena | Project/cruise | | "source" | provided by Tuerena |
| Tuerena | Longitude | $[0°, 360°]$ | $[-180°, 180°]$[3] | comparability |
| Lorrain | | | | |
| MacKenzie et al. (2019) | depth | original | deleted | suspected typo |
| Voss and von Bodungen (2003) | trap duration | original | deleted | suspected typo |
| De Jonge et al. (2015a) | Method | multiple investigations (MULT) | in-situ pump | found in De Jonge et al. (2015b) |

[1] By arithmetic mean.
[2] Only for sample durations entirely within an explicit month and year.
[3] We applied Equation 2.

The reference includes the citations as detailed as possible. Wherever available, this is taken from the original source. Otherwise, we tried to include author, title, publication year and platform and doi. For unpublished data like Harrison's from the Goericke's data set or those included by the coauthors, we denoted, from where we took the data.

Temporal and geographical locations determine sample location and time. Coordinates are given in decimal over $[-90°, 90°] \times [-180°, 180°]$. The sample depth is given in meters measured positively from the ocean surface downwards. Month and year are used to describe the sample date, specific days are neglected.

Anomalies of $\delta^{13}C_{POC}$ are given in the $\delta$-ratio described in Equation 1. A sample method was added, wherever available. Any special sampling circumstances are given in the "Note" column. Activity duration of sediment traps is denoted in the last column.

The "Origin" columns lists the associated project or cruise or author's note. Some samples were given with multiple project connections, all of them are given in this column.

### 3.2 Interpolated data sets

The interpolated $\delta^{13}C_{POC}$ data are available as Network Common Data Form (NetCDF) files on two global grids with different resolutions. NetCDF files are machine-independent and support creation, accessing and sharing of array-oriented scientific data.

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021


Open Access Earth System Science Data Discussions

**Table 3.** Available data and meta information: The columns of the raw data set correspond to the provided data and meta information. Their names are given in the first column of this table. The second holds a short description of their content, the third their ranges of values. In the final column we give how well this data kind is covered relative to the size of the full data set.

| column | content | range of values | coverage[1] |
|---|---|---|---|
| Reference | citation[2] | description | full[3] |
| No | running index | $\{1, ..., 4732\}$ | full |
| Lat | latitude in decimal[4] | $[-90°, 90°]$ | 4604 / 4732 |
| Long | longitude in decimal[4] | $[-180°, 180°]$ | 4604 / 4732 |
| d13Canomaly | $\delta^{13}C_{POC} - mean_{\delta^{13}C_{POC}}$[4] | $[-31.194384721885687, 19.455615278114315]$ | full |
| Temp | temperature in °Celsius[4] | $[-1.8, 31.12]$ | 1622 / 4732 |
| Month | month as number | $\{1, ..., 12\}$ | 4114 / 4732 |
| Year | years A.D. | $\{1964, ..., 2015\}$ | 4483 / 4732 |
| Depth | depth in m | $[0, 4850]$ | 3917 / 4754 |
| Method | measurement method of $\delta^{13}C_{POC}$ | description | 3164 / 4732 |
| Origin | associated project or cruise | description | 3921 / 4732 |
| Note | special circumstances, if | description | 140 / 4732 |
| Trap duration | duration of trap activity in days | $[1, 133]$ | 533 / 587[5] |

[1] Ratio of available entries relative to the full number of data points.
[2] Wherever possible, this includes: author(s), year, title, journal name, full, number, issue, pages and doi.
[3] Primary source was not available in every case as a reference. A note, where the data was taken is included in this case.
[4] With as many decimal places as available.
[5] Here, abundance is given relative to the full number of sediment trap samples.

On the coarser grid, we provide seven different files, where six of them each contain data of an individual decade (from the 1960s through the 2010s). The seventh file comprises a combined set of all interpolated $\delta^{13}C_{POC}$ data. On the finer grid, we provide one file including all $\delta^{13}C_{POC}$ measurements with complete spatial-temporal information.

For the coarser interpolation, we chose the grid of the version 2.9 UVic model, as used e.g. in Schmittner and Somes (2016), because this model includes simulation of $\delta^{13}C_{POC}$. The horizontal spatial grid consists of $100 \times 100$ cells with a resolution of $1.8° \times 3.6°$, arranged from 0 to 360° in longitude and $-90$ to 90° in latitude. The vertical grid is split up into 19 vertical layers, decreasing in resolution with depth. The two uppermost layers reach down to depths of 50 and 130 m respectively, which represent the upper ocean's euphotic zone.

The finer interpolation was carried out on the $1° \times 1°$ grid of the World Ocean Atlas (Garcia et al., 2010) and only includes $\delta^{13}C_{POC}$ data with full spatio-temporal metadata coverage. Here, the provided NetCDF file includes also the year ranging from 1964 to 2015 on the l/t axis and the month ranging from 1 to 12 on the m/e axis.

FERRET scripts were used for the interpolations. The interpolation function SCAT2GRIDGAUSS FERRET function by NOAA's Pacific Marine Environmental Laboratory performed the spatial averaging under PyFerret v7.5. Calculations in this

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021


Open Access Earth System Science Data Discussions

functions are based on a work by Kessler and McCreary (1992) and can be summarized as follows: Let $(x_1, y_1), ..., (x_n, y_n) \subseteq \mathbb{R}^2$ be an equidistant grid and $(\tilde{x_1}, \tilde{y_1}), ..., (\tilde{x_m}, \tilde{y_m}) \subseteq \mathbb{R}^2$ be irregular measurement locations of a real tracer $D_j, j \in \{1, ..., m\}$. Then the value $D_i \in \mathbb{R}$ at grid point $(x_i, y_i), i \in \{1, ..., n\}$ becomes interpolated as

$$D_i := \frac{\sum_{j=1}^{m} D_j W_{i,j}}{\sum_{j=1}^{m} W_{i,j}} \tag{4}$$

where

$$W_{i,j} := \begin{cases} 0; & \tau_{i,j} < e^{-CX} \\ 0; & \tau_{i,j} < e^{-CY} \\ \tau_{i,j}; & else \end{cases} \tag{5}$$

with $\tau_{i,j} := \exp\left(-\left(\frac{(x_j - x_i)^2}{X^2} + \frac{(y_j - y_i)^2}{Y^2}\right)\right)$ is the Gaussian weight function and $X, Y \in \mathbb{R}$ are scaling arguments and $C \in \mathbb{R}$ the cut-off parameter. We set to $X = 1.8$, $Y = 0.9$ and $C = 1$ in the our script.

Since the interpolation into the finer grid excluded all data without full spatio-temporal metadata coverage, we focus following descriptions of interpolated data on the coarse grid interpolations.

## 4 Main dataset characteristics

The final data set includes 4732 individual $\delta^{13}C_{POC}$ measurements of seawater samples. We show the distribution of $\delta^{13}C_{POC}$ values by Gaussian kernel density estimation (KDE). KDEs are a non-parametric density estimation (Silverman, 1986) for approximation of probability density functions, which is theoretically similar to a histogram but with a continuous curve not dependent on rigid intervals. We applied a Python implementation from the SciPy stats-package (Virtanen et al., 2020) to create the results presented here. Likewise, we derived conditional probability densities of $\delta^{13}C_{POC}$ values, given the different measurement method applied.

### 4.1 Range and outlier values

The data distribution is presented by its KDE in Figure 1. The interval of $\delta^{13}C_{POC}$ values ranges over $[-55.15, -4.5]$ with a mostly smooth distribution. Most of our data exhibit values around $\delta^{13}C_{POC} \approx -24‰$, which becomes clearly identifiable as a single maximum in the KDE. Two smaller modes are visible at around $\delta^{13}C_{POC} \approx -27.5‰$ and $\delta^{13}C_{POC} \approx -22‰$. A steep decline to zero follows after the two outer modes. The steep decline of the KDE stops at around $\delta^{13}C_{POC} = -37‰$ and $\delta^{13}C_{POC} \approx -14‰$. Between $\delta^{13}C_{POC} \approx -37‰$ and $\delta^{13}C_{POC} \approx -55.15‰$ as well as between $\delta^{13}C_{POC} \approx -14‰$ and $\delta^{13}C_{POC} \approx -4.5‰$ the KDE closely aligns to the x-axis, what indicates very little data points lying in this range.

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021


Open Access Earth System Science Data Discussions

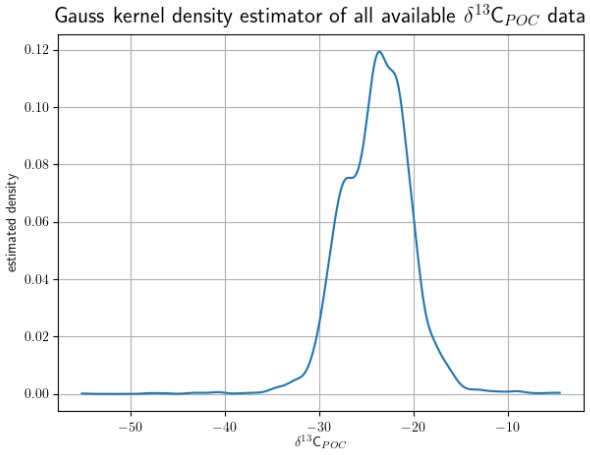


**Figure 1.** The density function of all individual $\delta^{13}C_{POC}$ measurements approximated by Gaussian kernel density estimation: Values of the estimated density are drawn on the y-axis, the $\delta^{13}C_{POC}$ values run on the x-axis. The higher the value of the estimated density is, the more $\delta^{13}C_{POC}$ points have been measured around this value.

Below $\delta^{13}C_{POC} = -37‰$ we find 17 data points ranging down to $\delta^{13}C_{POC} = -55.15‰$. Down to $\delta^{13}C_{POC} = -48‰$ these were all taken from Lein and Ivanov (2009) and Lein et al. (2006), measured in September or October 2003 and around the location $10°$ N, $104°$ W. The smallest outlier at $\delta^{13}C_{POC} = -55.15‰$ is taken from Altabet and Francois (2003an) from November 1996 and at $62.52°$ S, $169.99°$ E.

Above $\delta^{13}C_{POC} = -10‰$ we find 15 data points ranging up to $\delta^{13}C_{POC} = -4.5‰$. Three of them are taken from Lein et al. (2007) and measured at 800 $\mathrm{m}$ depth at a hydrothermal vent located $30.125°$ N, $42.117°$ W. Ten are taken from Calvert and Soon (2013b, c, a). All of these were measured between 636 and 901 $\mathrm{m}$ depth around $49°$ N, $130°$ W and all of them in February or May, but one in August. The final two are part of the Lorrain data set. Both were measured at the ocean surface, in July at $5.3°$ S, $164.9°$ E, and December at $20.9°$ S, $159.6°$ E.

Since more than 98 $\%$ of the data (4668 of the 4732 data points) have values that lie between $\delta^{13}C_{POC} = -35‰$ and $\delta^{13}C_{POC} = -15‰$, we will focus on this range in our following analyses.

We tested the robustness of our KDE approach in a subsampling experiment. We considered 500 random subsets of 20 $\%$ of the original data over the range with the highest data density $[-35, -15]$ and visualized their KDEs in Figure 2. They show peaks at $\delta^{13}C_{POC} \approx -23‰$ fitting the maximum and the second smaller mode right from it, and at $\delta^{13}C_{POC} \approx -27.5‰$. Outside $[-27, -22]$ the KDEs are closely aligned. Mean and standard variation of the KDE ensemble also shows the highest variability around the two modes at $\delta^{13}C_{POC} = -23‰$ and $\delta^{13}C_{POC} = -27.5‰$.

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021


Open Access Earth System Science Data Discussions

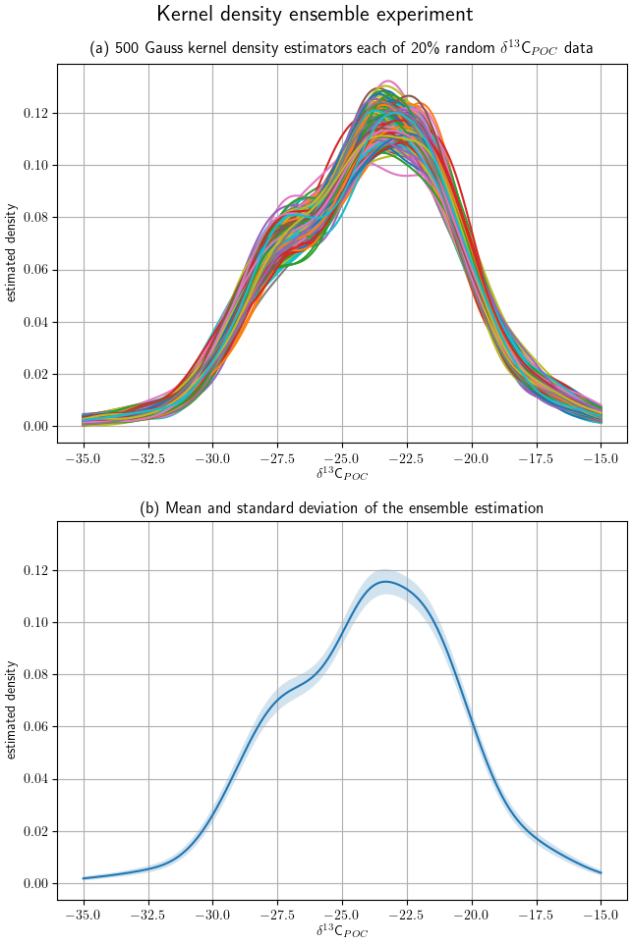


**Figure 2.** A random sample of 20% of the $\delta^{13}C_{POC}$ data was taken from the full data set for 500 times to generate an ensemble of subsets. Their densities were approximated with a Gaussian kernel density estimator. (a) shows all 500 estimated densities by individual lines. (b) shows the mean and the variance of the full ensemble of densities by a graph and the shaded area around it, respectively.

### 4.2 Sampling methods

Various sampling methods were involved in obtaining the $\delta^{13}C_{POC}$ data. We identified eighteen different sampling methods that could be attributed to 67 % of the data as meta information. In principle, all eighteen methods can be grouped into five main observational types: bottles, intake, nets, traps and diverse. Bottles include samples taken from Niskin bottles, PEP bottles and samples collected via Seabird submersible pumps. By "intake" we refer to all versions of pumps, underway cruise track measurements, as well as Multiple Unit Large Volume Filtration System (MULVFS). Nets represent all occurring versions of plankton nets and traps all represented sediment traps and moorings. Finally, the deep sea manned submersible (MIR2) is not classified to any of these groups and was assigned to a cluster that we refer to as "diverse".

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021




All sample devices provide data over all sample depths. Deeper samples were mainly taken from traps and pump systems, the upper from bottle and net data. Most data sampled deeper than 2600 m was collected by sediment traps. At 3800 m there were several trap contributions by Calvert (e.g. Calvert, 2002), mostly from the late 1980s. Data sampled by a deep-sea manned submersible is given at locations down to 2520 m (Lein and Ivanov (2009)).

We resolved differences between sampling methods in the Atlantic Ocean by comparing the KDE of all $\delta^{13}C_{POC}$ data with conditional probability densities of the same data distinguished by the four major methods in Figure 3. The Atlantic Ocean covers in this context the area between 45° S and 80° N and 70° W and 20° E. Overall, after accounting for spatial sampling bias by comparing with regions, the different methods are generally consistent with each other (Figure 3).

In the full Atlantic Ocean, densities of intake and net data are most representative of the maximum full $\delta^{13}C_{POC}$ sample. From the intake data shown here, $\approx 80$ % were sampled within 30° S and 30° N. When restricting to this area, net data resembles the full data better. But other than the intake data, of the net samples were $\approx 80$ % collected between 30° N and 60° N, where it also fits the overall $\delta^{13}C_{POC}$ density best, followed by trap data. Trap and bottle data deliver lowest $\delta^{13}C_{POC}$ measurements in the Atlantic Ocean. Both data kinds were with $\approx 74$ to 85 % sampled north from 60° N. A restriction to this area shows trap and bottle samples being close aligned to the full data in this region.

The variance of the intake and trap data is with $\approx 3‰$ a bit lower than the variance of all $\delta^{13}C_{POC}$ together, which is with $\approx 5‰$ the highest here presented. Bottle and net data both show a variance less than 2‰. Furthermore, trap, net and full $\delta^{13}C_{POC}$ show a clearly pronounced second second mode in their densities, while bottle and net data show a mostly clear individual maximum.

## 5 Spatial distribution

We show how the measurements are distributed over the ocean depths and surface. Most $\delta^{13}C_{POC}$ data has been measured in the uppermost few ocean meters and best surface coverage is available for the Atlantic Ocean. Changes in $\delta^{13}C_{POC}$ on the ocean surface were evaluated based at the coarse resolution gridded NetCDF data.

### 5.1 Vertical distribution of the data set

Depth values are available for more than 80 % of the sample data locating most of them in the upper ocean. This makes depth one of the least well covered metadata after temperature and sample method. The distribution of depth values is shown in Table 4 and an approximation by Gaussian KDE visualized in Figure 4. The KDE resolves best data coverage for the uppermost $\approx 500$ m of the oceans and a second far smaller maximum at $\approx 3800$ m. The depths ranges presented in Table 4 correspond to the depth intervals of the coarse interpolation, only the two uppermost layers are presented in more detail and the last four are combined. Within these first 130 m we observe highest data density and find nearly 2500 measurements of $\delta^{13}C_{POC}$, where already nearly 1000 of them were measured within $[0\,\mathrm{m}, 10\,\mathrm{m})$. Even in $[3430\,\mathrm{m}, 3900\,\mathrm{m})$ still 200 $\delta^{13}C_{POC}$ values are available. The two deepest values were taken from the Fischer (1989) and Altabet and Francois (2003av) and sampled at 4500 and 4850 m depth.

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021


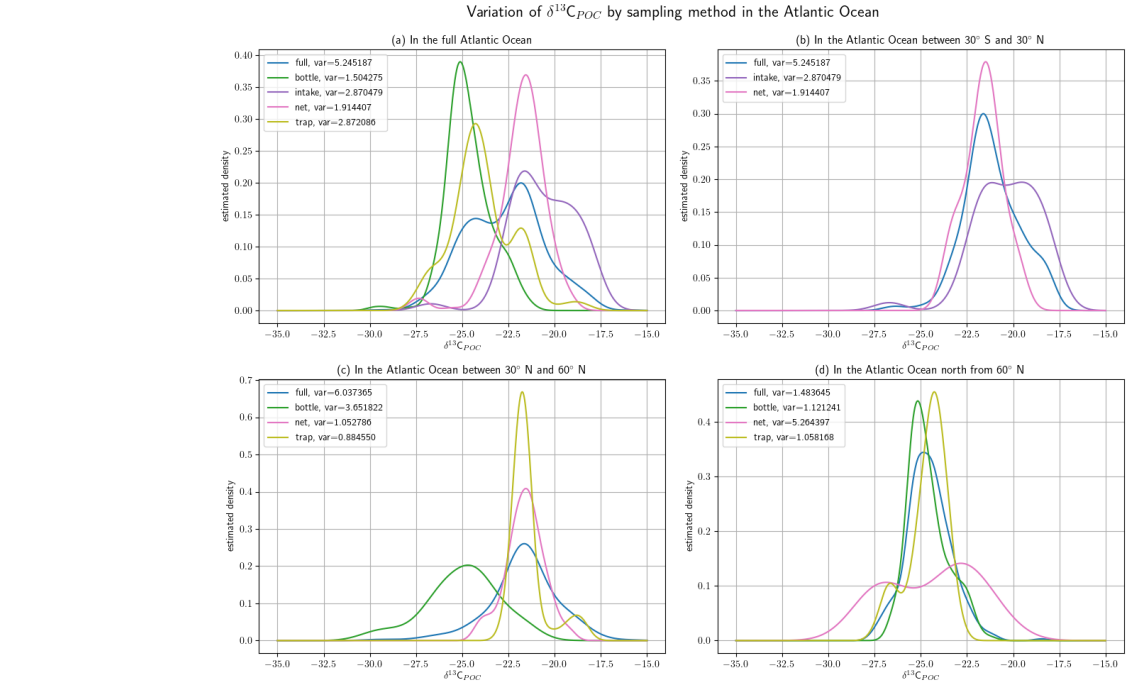


**Figure 3.** Separation of $\delta^{13}C_{POC}$ in the Atlantic Ocean data by four main sample methods: bottle, intake net and trap data. (a) shows the full Atlantic Ocean, (b) the equatorial core of the Atlantic Ocean, (c) the Atlantic between 30° S and 30° N and (d) its most northern area. In each plot, the density of the $\delta^{13}C_{POC}$ sample groups with enough data was approximated by Gaussian kernel density estimators and drawn with an individual color. An additional graph shows the comparison to the full $\delta^{13}C_{POC}$ data density in the respective area.

### 5.2 Horizontal distribution of the data set

All global oceans are covered with $\delta^{13}C_{POC}$ data. In Figure 5 the horizontal distribution of available data is depicted for both interpolations. Here, the coarse resolution interpolation is independent of time and the fine resolution averaged over all included times. A similar plot, but with a different purpose, is given later in this work in Figure 9 showing only surface data locations.

Many cruises are visible as lines formed by connected grid cells in Figure 5. Especially in the Atlantic and Indian Ocean and shorter in the Southern Ocean. Also, data locations of smaller individual or connected grid cells occur. These are mainly located in the Pacific, Arctic and Southern Ocean. The Atlantic Ocean provides best data coverage. Following, the Southern and Indian oceans contain the next best coverage with the northern Pacific having the sparsest.

Highest $\delta^{13}C_{POC}$ values are evident in low latitude regions reaching down to $\leq -30‰$. In the Atlantic Ocean highest values are measured between 0-30° N and 30-60° W as well as close to the western coast of France reaching up to $\geq -17‰$. The Indian Ocean shows generally high values with $\approx -20‰$. In the Pacific Ocean highest values are close to the Peruvian coast

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021


Open Access Earth System Science Data Discussions

**Table 4.** Vertical data coverage in depth layers inspired by the coarse interpolation grid: The first column lists the observed depth layers. Below 50 m they are as defined by the coarse grid used for interpolation of the $\delta^{13}C_{POC}$ data. The second columns gives the explicit number of $\delta^{13}C_{POC}$ data points available in this depths range.

| depth range | $\delta^{13}C_{POC}$ values available |
|---|---|
| $[0\,m, 1\,m)$ | 329 |
| $[1\,m, 5\,m)$ | 351 |
| $[5\,m, 10\,m)$ | 286 |
| $[10\,m, 25\,m)$ | 480 |
| $[25\,m, 50\,m)$ | 319 |
| $[50\,m, 130\,m)$ | 693 |
| $[130\,m, 240\,m)$ | 351 |
| $[240\,m, 380\,m)$ | 147 |
| $[380\,m, 550\,m)$ | 174 |
| $[550\,m, 750\,m)$ | 64 |
| $[750\,m, 980\,m)$ | 131 |
| $[980\,m, 1240\,m)$ | 75 |
| $[1240\,m, 1530\,m)$ | 26 |
| $[1530\,m, 1850\,m)$ | 7 |
| $[1850\,m, 2200\,m)$ | 50 |
| $[2200\,m, 2580\,m)$ | 131 |
| $[2580\,m, 2990\,m)$ | 66 |
| $[2990\,m, 3430\,m)$ | 35 |
| $[3430\,m, 3900\,m)$ | 200 |
| $[3900\,m, 6080\,m)$ | 2 |

and Papua New Guinea. We also find high values in the Bering Strait and and the northern edge of the Southern Ocean around 65° E.

Lowest $\delta^{13}C_{POC}$ values are mostly found in the Southern Ocean. Nearly all measured grid cells here belong to $\delta^{13}C_{POC}$ values lower than around $-28‰$. The Arctic Ocean shows low values as well, for instance in the Kara Sea. Lowest values in the Pacific Ocean occur in the Southern Ocean at high latitudes.

### 5.3 Meridional trend of $\delta^{13}C_{POC}$ values

We show the north-south trend of $\delta^{13}C_{POC}$ over the Atlantic Ocean based on the six coarse grid interpolations, each over one of the available decades 1960s to 1990s. We chose this section due to its best data coverage. A biome mask according

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021


Open Access Earth System Science Data Discussions

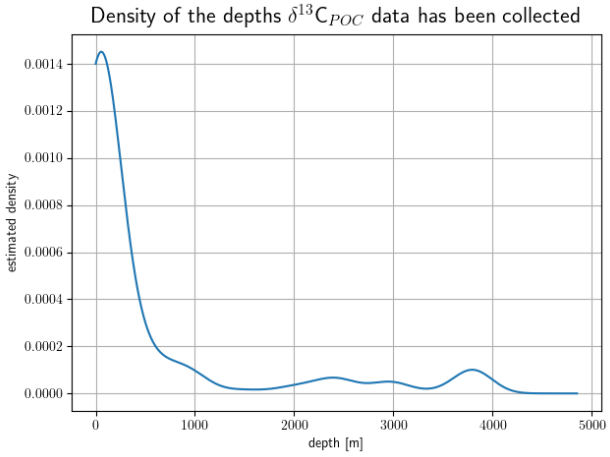


**Figure 4.** The vertical distribution of available $\delta^{13}C_{POC}$ samples is given by the approximated density of the measurement depths. The y-axis shows the estimated density of the depth values, the x-axis the depth in m. The estimation was realized by a Gaussian kernel density estimator. Its altitude correlates to the amount of $\delta^{13}C_{POC}$ data been measured at this depth.

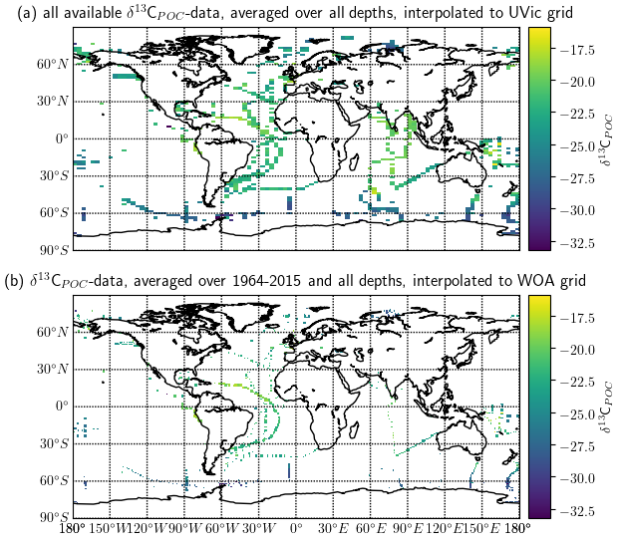


**Figure 5.** Global distribution of the $\delta^{13}C_{POC}$ data is visualized based on its interpolation in (a) the UVic and (b) WOA grids. The data used for (a) are independent of time and include all available measurements with spatial information. The data shown in (b) includes only data with complete spatio-temporal metadata and are averaged over time. The data is averaged over all available depth levels. Each colored square refers to a grid cell with available $\delta^{13}C_{POC}$ measurements. The colors indicate the $\delta^{13}C_{POC}$ value in the respective grid cell.

to Fay and McKinley (2014) was applied to the gridded data, thereby defining latitudinal zones in the entire Atlantic Ocean.

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021


Open Access Earth System Science Data Discussions

**Table 5.** Data coverage within the available decades: The first column lists the available decades, the second column the number of sampled $\delta^{13}C_{POC}$ data points within this time frame.

| decade | $\delta^{13}C_{POC}$ values available |
|---|---|
| 1960s | 74 |
| 1970s | 321 |
| 1980s | 463 |
| 1990s | 2403 |
| 2000s | 614 |
| 2010s | 589 |

Distributions of $\delta^{13}C_{POC}$ within the biomes are shown in Figure 6. Different colors mark the individual biomes and a black line shows the general global $\delta^{13}C_{POC}$ distribution.

The biomes derived by Fay and McKinley (2014) are areas with consistent biological and ecological properties. The chosen biomes cover the Atlantic Ocean and extend to the Arctic Sea and parts of the Southern Ocean. Laterally the biomes are restricted to 70° W and 20° E. We adopted the numbering of the biomes from Fay and McKinley (2014), 9 to 17, but 14, where 15 to 17 had to be cut to the given lateral range. Their location in the Atlantic Ocean is also shown in Figure 10.

Observations by the biomes are consistent with the ones from Figure 5. The two biomes showing the lowest $\delta^{13}C_{POC}$ values from -28 to -29 /permil are those two located farthest south. The biome located farthest north contain the next lowest value at about -24 /permil. The final biomes are in the lower latitudes show similarly higher $\delta^{13}C_{POC}$ values from -23 to $-21‰$ .

## 6 Temporal distribution of the dataset

The full $\delta^{13}C_{POC}$ data cover a time period of around 50 years over 1964-2015 and all twelve months. The number of samples measured during individual decades varies considerably with most measurements in the 1990s. Coverage within the months is quite comparable, only the months January, March, July, and December exhibit less data.

The distribution of $\delta^{13}C_{POC}$ samples over the years is resolved in Table 5 and visually approximated by Gaussian KDE in Figure 5. The 1990s show best data coverage. More than half of the data points is associated to a year in this decade, which is visible by a pronounced maximum in the estimated density. Sparsest data is found in the 1960s, where only 74 data points were sampled. All other decades come with between around 300 and 600 $\delta^{13}C_{POC}$ data points. The latest data is mostly from Lorrain, MacKenzie et al. (2019) and Kaiser et al. (2019). The oldest data is taken from the data sets by Tuerene, Degens et al. (1968) and Eadie and Jeffrey (1973).

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021


Open Access Earth System Science Data Discussions

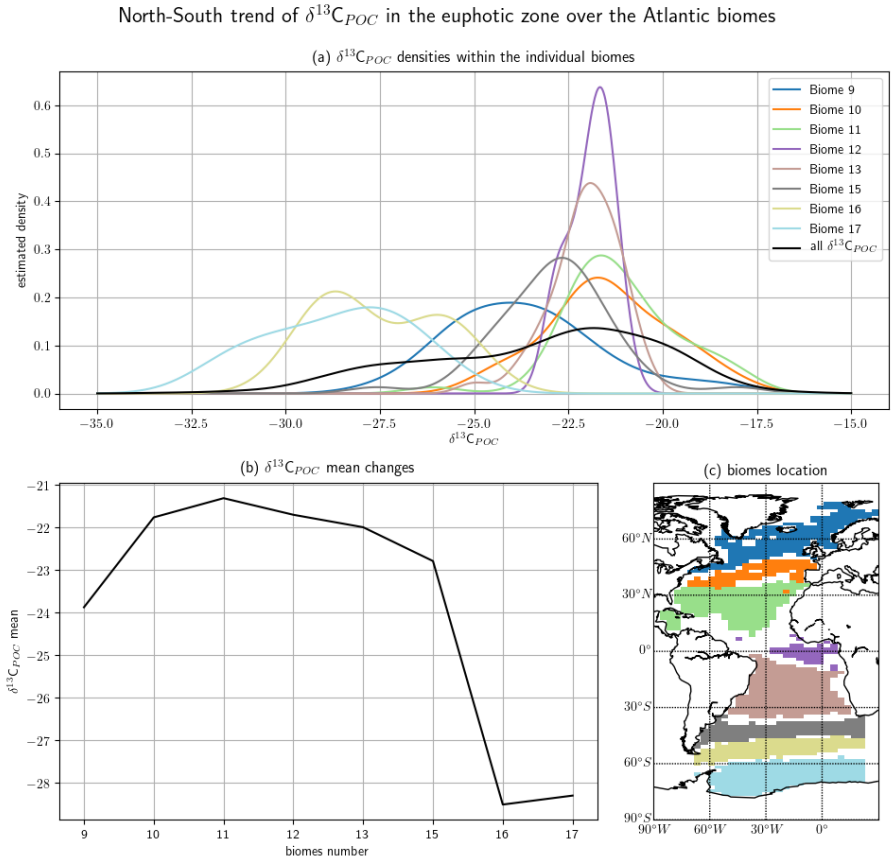


**Figure 6.** North-south trend of sampled $\delta^{13}C_{POC}$ values is visualized by a cross section over the Atlantic ocean. Biomes define the latitudinal bands of the interpolated data set. (a) For each biome a Gaussian kernel density estimator approximates the density of the contained $\delta^{13}C_{POC}$ data. All of the densities are drawn together, each color indicates the respective biome, a final graph the estimated density of all $\delta^{13}C_{POC}$ data. (b) A mean-vs-biome plot shows the steep decline from the tropical biomes towards the higher latitudes. (c) Each biome is drawn in the color of its corresponding density estimate in (a) above. The biome numbers increase from the north to the south.

### 6.1 Seasonal trends

Monthly clustered data of northern and southern hemisphere show seasonal variations, but more observations are required to
300 demonstrate robust seasonality within different regions. Since most of the available $\delta^{13}C_{POC}$ data originates in the 1990s, we selected only data from this decade to exclude changes that might be introduced by longer term changes. In Figure 8 we displayed all months with enough data points for construction of a comparable KDE and indicate same months by same colors. The KDEs are supported by comparison of the median values of the individual months.

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021


Open Access Earth System Science Data Discussions

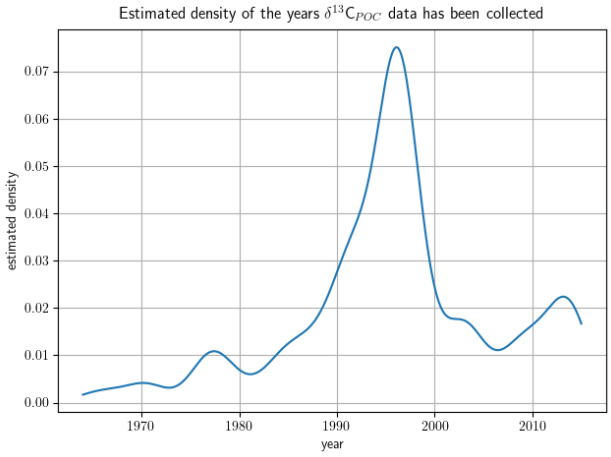


**Figure 7.** The distribution of $\delta^{13}C_{POC}$ data samples over the years approximated by Gaussian kernel density estimation. On the y-axis the density is drawn, on the x-axis the sample year. Higher altitude of the graph indicates years with more available data.

Northern hemisphere $\delta^{13}C_{POC}$ values are generally higher than in the southern hemisphere. Highest values in the northern hemisphere can be found in April, September and October. They all show a maximum around $\delta^{13}C_{POC} \approx -21‰$. In the southern hemisphere these months together with November show the lowest values around $\delta^{13}C_{POC} \approx -28‰$. Winter months in both hemispheres show middle $\delta^{13}C_{POC}$ values centering on the northern around $\delta^{13}C_{POC} \approx -25‰$ and the southern hemisphere around $\delta^{13}C_{POC} \approx -27‰$. Summer data are far better available on the northern hemisphere, where these data are also close to the winter data. Lowest but few data in the northern hemisphere are found in June, May, September and February. A strong pronounced highest maximum in the southern hemisphere can be seen in March.

### 6.2 Multi-decadal trends

The decadal interpolations into the coarse grid are basis for showing long term changes in the $\delta^{13}C_{POC}$ data. An overview of where the data within the individual decades were sampled is given in Figure 9. This shows that sparsest coverage is obtained in the 1960s. Half of the data are sampled in the Southern Ocean between 70° W and 20° E, the other half closely located to central American continent. Most data in the Indian Ocean were sampled in the 1970s. A cruise across the southern part of the Atlantic Ocean up to 30° N and some samples close to Iceland were also measured in this decade. 1980s are similarly sparse as the 1960s. Measurements of the 1980s were taken at locations similar to the samples collected in the 1960s in the Southern Ocean and in the Arctic and the Atlantic close to the equator. 1990s are best covering most ocean basins. Most Southern Ocean data is sampled within them. The 2000s provide a good coverage of the Atlantic and the Arctic Ocean. Finally, the 2010s data were mostly sampled in the southern hemisphere in the open Pacific and Atlantic. Some smaller Eurasian continental sea data are also part of the 2010s samples.

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021


Open Access Earth System Science Data Discussions

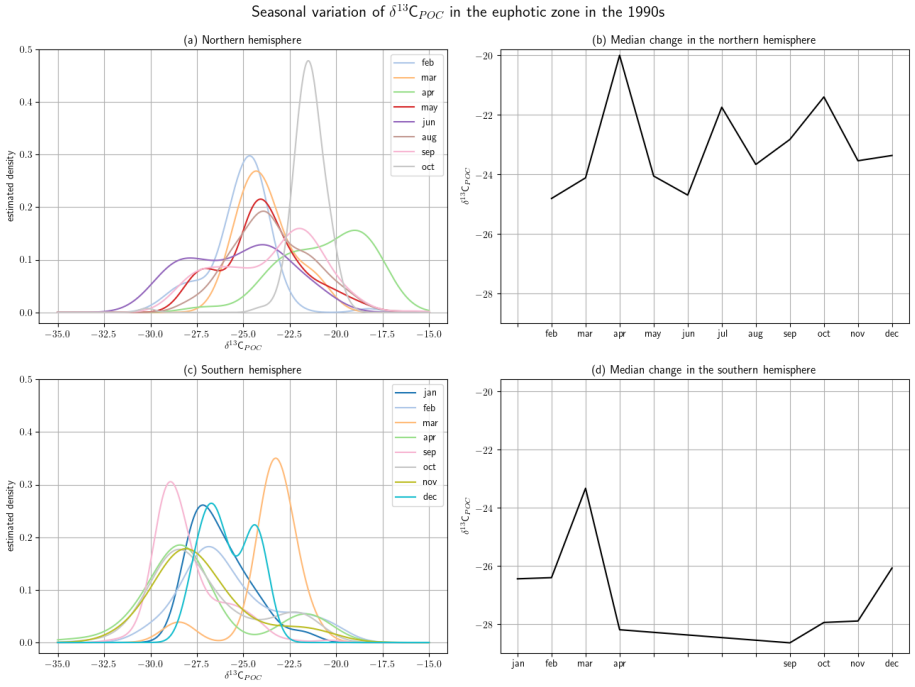


**Figure 8.** Seasonal variations are split up by hemisphere in the northern in (a + b) and southern in (c + d). Due to their best data coverage, the analyses are carried out within the 1990s. The $\delta^{13}C_{POC}$ is split up by sample month and for every month with enough available data points a Gaussian kernel density estimator approximate their density in (a) and (c). Not all months include enough data for a density estimation. For each hemisphere the densities are drawn all together, each month indicated by an individual color. The right column shows a month-vs-mean plot for the (b) northern and (d) southern hemisphere. The x-axis has labels for every month, where data for a mean calculation was available.

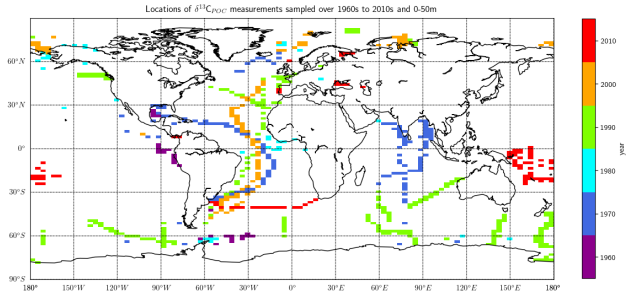


**Figure 9.** Sample locations of the $\delta^{13}C_{POC}$ data are marked regarding to the interpolation grid. Only data of the uppermost layer are considered in this plot. The different colors indicate the different sample decades.

We show the changes in $\delta^{13}C_{POC}$ values over the available decades in Figure 10. The plot includes approximated densities of the $\delta^{13}C_{POC}$ measurements for each decade and median-vs-years graphs. The Southern Ocean was excluded from the main analysis due to the sparse coverage outside of the 1990s and showed its few available results in the two lower separate panels.

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021


Open Access Earth System Science Data Discussions

The Southern Ocean is defined here as the ocean area south of 45° S. All these analyses are restricted to the euphotic zone, i.e. the uppermost 130 m resembling the two first layers of the grid.

A clear decrease in $\delta^{13}C_{POC}$ values can be identified for the global ocean outside of the Southern Ocean. All, but the 1960s and 1980s show one clear maximum in their approximated densities. The two exceptions are the decades that both show a second expressed second density maximum at lower values. The main maxima shift from the 1960s at $\delta^{13}C_{POC} \approx -19.9‰$ with every decade lower to the 2010s at $\delta^{13}C_{POC} \approx -23‰$. This decrease is also clearly visible in the comparison of the decadal means.

The Southern Ocean provides far worse data coverage. Only the 1960s, 1980s and 1990s include enough data to construct a comparable KDE. A median could also be calculated from the 1970s and 2000s. But due to this very little available data, all of this results must be seen with highest caution. Except for the 1980s the Southern Ocean shows an increase in $\delta^{13}C_{POC}$ values over the decades. The 1960s show a median and main maximum at $\delta^{13}C_{POC} \approx -27.5‰$. The median increases to the 2000s to $\delta^{13}C_{POC} \approx -20‰$ with a pronounced dip in the 1980s down to less than $\delta^{13}C_{POC} \approx -30‰$. The densities support this observation in the 1980s, where the maximum is below $\delta^{13}C_{POC} \approx -30‰$. Nevertheless, we need to take into account that most Southern Ocean data were sampled in the 1990s, while the 1970s and 2000s provide only few data and might not deliver comparable results.

## 7 Conclusions

The aim of this work was to construct the largest publicly accessible $\delta^{13}C_{POC}$ data set. We tackled this by merging all known data sets and collecting all available additional seawater samples from a free data distribution platform (PANGAEA). This newly constructed $\delta^{13}C_{POC}$ data set currently contains 4732 data points with the potential to grow in the future. It is provided in a csv structure and interpolated on two different resolution global grids as NetCDF format. The csv file contains the $\delta^{13}C_{POC}$ with respect to their mean and all available meta information. The interpolations are provided on a coarse $1.8° \times 3.6°$ grid of a $\delta^{13}C_{POC}$ simulating model and a finer $1° \times 1°$ grid by the World Ocean Atlas. We provided a detailed description of our data collection procedure, all added meta information and their coverage as well of the interpolation procedure carried out. We took highest care to make all data coherent, comparable and back trackable and all adjustments transparent. Assumptions, changes and deletions of the used data sets are described in detail.

We described the general spatial and temporal trends of the sampled $\delta^{13}C_{POC}$ data by the raw data file. Distributions were always approximated by Gaussian kernel density estimators. The data ranges from 1964 - 2015 with far best coverage in the 1990s. Sample locations reach down to a depth of nearly 5000 m and best covers the uppermost 10 m in some areas. We were able to show our $\delta^{13}C_{POC}$ data values are mostly located between $\delta^{13}C_{POC} = -15‰$ and $\delta^{13}C_{POC} = -35‰$ with two maxima at around $\delta^{13}C_{POC} = -27‰$ and $\delta^{13}C_{POC} = -23‰$, the latter one being the more pronounced. A comparison of the main sample methods showed consistent results when compared with regions. $\delta^{13}C_{POC}$ data separated by months indicate counteracting seasonal trends on both hemispheres, but more data is required to demonstrate robust seasonality.

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021


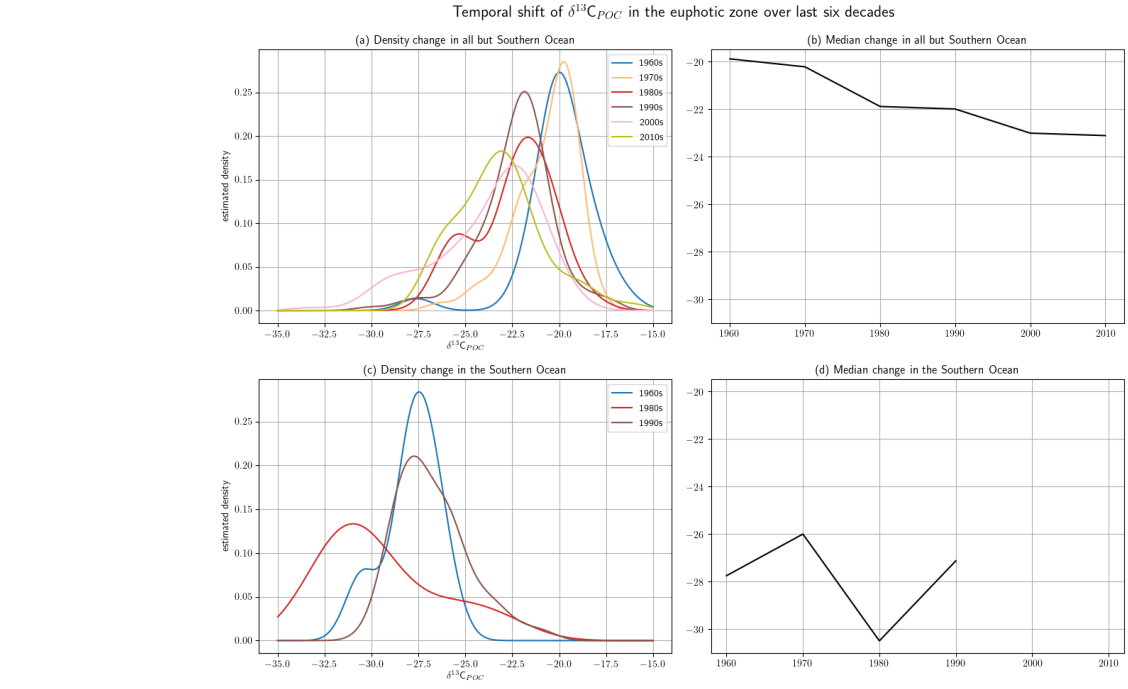


**Figure 10.** The decadal shift of $\delta^{13}C_{POC}$ values for all, but the Southern Ocean (a) and (b) in the upper and only the Southern Ocean (c) and (d): (a) and (c) show estimated densities of $\delta^{13}C_{POC}$ values. The differently colored graphs refer to the individual decades. (b) and (d) show the changes of the $\delta^{13}C_{POC}$ decadal mean against the decades. Southern Ocean data is sparsely covered in the 1970s, not delivering enough results for a kernel density estimate. It is not sampled in the 2010s and only with a single value in the 2000s, hence these are excluded.

The interpolated data provide insights in geographical behavior of the sampled $\delta^{13}C_{POC}$ data. We showed a good general coverage of all global oceans by $\delta^{13}C_{POC}$ data although the northern Pacific is still sparsely covered. Since the Atlantic Ocean provides the best coverage, corresponding data were used for a north-south trend analysis. This supported the observation that lowest values ($<\approx -28‰$) can be found in the Southern Ocean whereas highest ($>\approx -22‰$) are restricted to low latitudinal regions. Finally, we showed the sample locations and value development of $\delta^{13}C_{POC}$ over the observed decades. Since the Southern Ocean data was mainly sampled in the 1990s, a significant multi-decadal trend could not be detected there. In all other oceans our $\delta^{13}C_{POC}$ data show a decrease by about 3‰ over the observed timeframe, which is about the double rate of the known Suess effect (Keeling, 1979) on aqueous $\delta^{13}CO_2$ (Young et al., 2013). This corroborates an increase in phytoplankton carbon fractionation that may be associated with a change in phytoplankton communities as previously suggested (Lorrain et al., 2020; Young et al., 2013). The $\delta^{13}C_{POC}$ dataset shows promise to better understand, constrain and predict carbon cycling in the future.

*Data availability.* The described $\delta^{13}C_{POC}$ data by Verwega et al. (2021) are available at https://doi.org/10.1594/PANGAEA.929931

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021




*Author contributions.* M. - Th. Verwega collected and merged the data, conducted the coarse-grid interpolation, performed the analyses and set up the manuscript. C. J. Somes initiated and supported the data collection, provided the coarse-grid interpolation scripts, conducted the fine-grid interpolation, guided analyses of the data and structured and proofread the manuscript. M. Schartau supported the data collection, guided their analyses and proofread the manuscript. R. E. Tuerena provided additional data and ideas for its analyses and proofread the manuscript. A. Lorrain provided additional data and proofread the manuscript. A. Oschlies guided the analysis of the data and proofread the maniscript. Th. Slawig guided the elaboration of the manuscript, structured and proofread it.

*Competing interests.* The authors declare that they have no conflict of interes.

*Acknowledgements.* The first author is funded through the Helmholtz School for Marine Data Science (MarDATA), Grant No. HIDSS-0005.
C. Somes is funded by the Deutsche Forschungsgemeinschaft (DFG, project no. 445549720)
We like to thank Tronje Kemena for providing the basic global biomes masks, used for analyzing the interpolated data sets on the coarse grid.

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021




## References

AESOPS: U.S. JGOFS Antarctic Environment and Southern Ocean Process Study, http://usjgofs.whoi.edu/southern.html.

Alfred-Wegener-Institut: PANGAEA Data Publisher for Earth & Environmental Science, https://www.pangaea.de.

Altabet, M. A. and Francois, R.: Natural nitrogen and carbon stable isotopic composition in surface water at cruise NBP96-05, https://doi.org/10.1594/PANGAEA.128266, 2003an.

Altabet, M. A. and Francois, R.: Natural nitrogen and carbon stable isotopic composition of station NBP96-05-06-4, https://doi.org/10.1594/PANGAEA.128229, 2003av.

Bidigare, R. R., Fluegge, A., Freeman, K. H., Hanson, K. L., Hayes, J. M., Hollander, D., Jasper, J. P., King, L. L., Laws, E. A., Milder, J., Millero, F. J., Pancost, R., Popp, B. N., Steinberg, P. A., and Wakeham, S. G.: Consistent fractionation of13C in nature and in the laboratory: Growth-rate effects in some haptophyte algae, Global Biogeochemical Cycles, 11, 279–292, https://doi.org/10.1029/96gb03939, https://doi.org/10.1029/96gb03939, 1997.

Buchanan, P. J., Matear, R. J., Chase, Z., Phipps, S. J., and Bindoff, N. L.: Ocean carbon and nitrogen isotopes in CSIRO Mk3L-COAL version 1.0: a tool for palaeoceanographic research, Geoscientific Model Development, 12, 1491–1523, https://doi.org/10.5194/gmd-12-1491-2019, https://doi.org/10.5194/gmd-12-1491-2019, 2019.

Calvert, S. E.: Stable isotope data of sediment trap P84-4, https://doi.org/10.1594/PANGAEA.68555, 2002.

Calvert, S. E. and Soon, M.: Carbon and nitrogen data measured on water samples from the multiple unit large volume filtration system (MULVFS) during John P. Tully cruise IOS_96-09, https://doi.org/10.1594/PANGAEA.808319, 2013a.

Calvert, S. E. and Soon, M.: Carbon and nitrogen data measured on water samples from the multiple unit large volume filtration system (MULVFS) during John P. Tully cruise IOS_96-18, https://doi.org/10.1594/PANGAEA.808320, 2013b.

Calvert, S. E. and Soon, M.: Carbon and nitrogen data measured on water samples from the multiple unit large volume filtration system (MULVFS) during John P. Tully cruise IOS_97-02, https://doi.org/10.1594/PANGAEA.808321, 2013c.

Cassar, N., Laws, E. A., and Popp, B. N.: Carbon isotopic fractionation by the marine diatom Phaeodactylum tricornutum under nutrient- and light-limited growth conditions, Geochimica et Cosmochimica Acta, 70, 5323–5335, https://doi.org/10.1016/j.gca.2006.08.024, https://doi.org/10.1016/j.gca.2006.08.024, 2006.

Chang, A. S., Bertram, M. A., Ivanochko, T. S., Calvert, S. E., Dallimore, A., and Thomson, R. E.: (Supplement 2) Total mass flux, geochemistry and abundance of selected diatom taxa of Effingham Inlet OSU Trap samples, PANGAEA, https://doi.org/10.1594/PANGAEA.806329, in supplement to: Chang, AS et al. (2013): Annual record of particle fluxes, geochemistry and diatoms in Effingham Inlet, British Columbia, Canada, and the impact of the 1999 La Niña event. Marine Geology, 337, 20-34, https://doi.org/10.1016/j.margeo.2013.01.003, 2013.

De Jonge, C., Stadnitskaia, A., Hopmans, E. C., Cherkashov, G. A., Fedotov, A., Streletskaya, I., Vasiliev, A. A., and Sinninghe Damsté, J. S.: (Table 2) Particulate organic carbon contentand the stable carbon isotope signal of suspended particulate matter samples, PANGAEA, https://doi.org/10.1594/PANGAEA.877962, in supplement to: De Jonge, C et al. (2015): Drastic changes in the distribution of branched tetraether lipids in suspended matter and sediments from the Yenisei River and Kara Sea (Siberia): Implications for the use of brGDGT-based proxies in coastal marine sediments. Geochimica et Cosmochimica Acta, 165, 200-225, https://doi.org/10.1016/j.gca.2015.05.044, 2015a.

De Jonge, C., Stadnitskaia, A., Hopmans, E. C., Cherkashov, G. A., Fedotov, A., Streletskaya, I., Vasiliev, A. A., and Sinninghe Damsté, J. S.: Drastic changes in the distribution of branched tetraether lipids in suspended matter and sediments from the Yenisei River and Kara

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021




Sea (Siberia): Implications for the use of brGDGT-based proxies in coastal marine sediments., Geochimica et Cosmochimica Acta, pp. 200–225, https://doi.org/10.1016/j.gca.2015.05.044, 2015b.

Degens, E. T., Behrendt, M., Gotthardt, B., and Reppmann, E.: Metabolic fractionation of carbon isotopes in marine plankton - II. Data on samples collected off the coasts of Peru and Ecuador, Deep Sea Research and Oceanographic Abstracts, 15, 11–20, https://doi.org/10.1016/0011-7471(68)90025-9, 1968.

Eadie, B. J. and Jeffrey, L. M.: $\delta^{13}$C analyses of oceanic particulate matter, Marine Chemistry, 1, 199–209, https://doi.org/10.1016/0304-4203(73)90004-2, 1973.

EurOBIS Data Management Team: PANGAEA - data from Archive of Ocean Data, http://ipt.vliz.be/eurobis/resource?r=pangaea_2724.

Fay, A. R. and McKinley, G. A.: Global open-ocean biomes: mean and temporal variability, Earth System Science Data, 6, 273–284, https://doi.org/10.5194/essd-6-273-2014, https://doi.org/10.5194/essd-6-273-2014, 2014.

Fischer, G.: Stabile Kohlenstoff-Isotopen in partikulärer organischer Substanz aus dem Südpolarmeer (Atlantischer Sektor), Ph.D. thesis, Bremen University, 1989.

Fontugne, M. and Duplessy, J. C.: Carbon isotope ration of marine plankton related to surface water masses, Earth and Plametary Science Letters, 41, 365–371, https://doi.org/10.1016/0012-821X(78)90191-7, 1978.

Fontugne, M. and Duplessy, J. C.: Oceanic carbon isotopic fractionation by marine plankton in the temperature range of -1 to 31°C, Oceanologica Acta, 4, 85–90, 1981.

Fontugne, M., Descolas-Gros, C., and de Billy, G.: The dynamics of CO2 fixation in the Southern Ocean as indicated by carboxylase activities and organic carbon isotopic ratios, Marine Chemistry, 35, 371–380, https://doi.org/10.1016/S0304-4203(09)90029-9, 1991.

Francois, R., Atlabet, M. A., Goericke, R., McCorkle, D. C., Brunet, C., and Posson, A.: Changes in the $\delta^{13}$C of surface water particulate organic matter across the subtropical convergence in the SW Indian Ocean, Global Biogeochemical Cycles, 7, 627–644, https://doi.org/10.1029/93GB01277, 1993.

Freeman, K. H. and Hayes, J. M.: Fractionation of carbon isotopes by phytoplankton and estimates of ancient $CO_2$ levels, Global Biogeochemical Cycles, 6, 185–198, https://doi.org/10.1029/92GB00190, 1992.

Fry, B.: $^{13}$C/$^{12}$C fractionation by marine diatoms, Marine Ecology Progress Series, 134, 283–294, https://doi.org/10.3354/meps134283, https://doi.org/10.3354/meps134283, 1996.

Fry, B. and Sherr, E.: $\delta^{13}$C Measurements as Indicators of Carbon Flow in Marine and Freshwater Ecosystems, 1989.

Garcia, H. E., Locarnini, R. A., Boyer, T. P., Antonov, J. I., Zweng, M. M., Baranova, O. K., and Johnson, D. R.: Nutrients (phosphate, nitrate, silicate), World Ocean Atlas 2009, 4, 2010.

Goericke, R.: Variations of marine plankton $\delta^{13}$C with latitude, temperature, and dissolved $CO_2$ in the world ocean, Global Biogeochemical Cycles, 8, 85–90, https://doi.org/10.1029/93GB03272, 1994.

Harrison: Unpublished data, quoted from Goericke (1994).

Hayes, J. M.: An Introduction to Isotopic Calculations, http://www.whoi.edu/cms/files/jhayes/2005/9/IsoCalcs30Sept04_5183.pdf, 2004.

Hofmann, M., Wolf-Gladrow, D. A., Takahashi, T., Sutherland, S. C., Six, K. D., and Maier-Reimer, E.: Stable carbon isotope distribution of particulate organic matter in the ocean: a model study, Marine Chemistry, 72, 131–150, https://doi.org/10.1016/s0304-4203(00)00078-5, https://doi.org/10.1016/s0304-4203(00)00078-5, 2000.

IPCC: Summary for policymakers, pp. 3–29, Cambridge University Press, Cambridge, UK, https://doi.org/10.1017/CBO9781107415324.004, 2013.

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021




IPCC: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [ Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)], 2014.

Jahn, A., Lindsay, K., Giraud, X., Gruber, N., Otto-Bliesner, B. L., Liu, Z., and Brady, E. C.: Carbon isotopes in the ocean model of the Community Earth System Model (CESM1), Geoscientific Model Development, 8, 2419–2434, https://doi.org/10.5194/gmd-8-2419-2015, https://doi.org/10.5194/gmd-8-2419-2015, 2015.

Jasper, J. P. and Hayes, J. M.: A carbonisotopic record of $CO_2$ levels during the late Quaternary, Nature, pp. 462–464, https://doi.org/10.1038/347462a0, 1990.

JGOFS: Joint Global Ocean Flux Study, http://ijgofs.whoi.edu.

Kaiser, D., Konovalov, S. K., Arz, H. W., Voss, M., Krüger, S., Pollehne, F., Jeschek, J., and Waniek, J. J.: Black Sea water column dissolved nutrients and dissolved and particulate organic matter from winter 2013, Maria S. Merian cruise MSM33, https://doi.org/10.1594/PANGAEA.898717, supplement to: Kaiser, David; Konovalov, Sergey K; Schulz-Bull, Detlef; Waniek, Joanna J (2017): Organic matter along longitudinal and vertical gradients in the Black Sea. Deep Sea Research Part I: Oceanographic Research Papers, 129, 22-31, https://doi.org/10.1016/j.dsr.2017.09.006, 2019.

Keeling, C. D.: The Suess effect: 13Carbon-14Carbon interrelations, Environment International, 2, 229 – 300, https://doi.org/10.1016/0160-4120(79)90005-9, http://www.sciencedirect.com/science/article/pii/0160412079900059, 1979.

Kessler, W. S. and McCreary, J. P.: The annual wind-driven Rossby wave in the subthermocline equatorial Pacific, Journal of Physical Oceanography, pp. 1192–1207, 1992.

Laws, E. A., Popp, B. N., Bidigare, R. R., Kennicutt, M. C., and Macko, S. A.: Dependence of phytoplankton carbon isotopic composition on growth rate and [CO2]aq: Theoretical considerations and experimental results, Geochimica et Cosmochimica Acta, 59, 1131–1138, https://doi.org/10.1016/0016-7037(95)00030-4, https://doi.org/10.1016/0016-7037(95)00030-4, 1995.

Lein, A. Y. and Ivanov, M. V.: (Table 9.4.3) Concentrations of suspended matter in water samples from the 9°50'N EPR hydrothermal field and contents and isotopic compositions of organic carbon in suspended matter, PANGAEA, https://doi.org/10.1594/PANGAEA.771566, in supplement to: Lein, AY; Ivanov, MV (2009): Biogeokhimicheskii Tsikl Metana v Okeane (Biogeochemical Cycle of Methane in the Ocean). Nauka Publ. (Moscow); Lisitzin, A.P. (Ed.), 576 pp., 2009.

Lein, A. Y., Bogdanov, Y. A., Grichuk, D. V., Rusanov, I. I., and Sagalevich, A. M.: (Table 5) Concentration of particulate organic carbon and its isotopic composition in water samples from hydrothermal fields at the axis of the East Pacific Rise near 9°50'N, PANGAEA, https://doi.org/10.1594/PANGAEA.745910, in supplement to: Lein, AY et al. (2006): Geochemistry of hydrothermal solutions from 9°50'N at the East Pacific Rise twelve years after eruption of a submarine volcano. Geokhimiya (Geochemistry), No 7, 749-762, 2006.

Lein, A. Y., Bogdanova, O. Y., Bogdanov, Y. A., and Magazina, L. O.: (Table 6) Isotopic composition of organic carbon from microbial communities within the Lost City hydrothermal field, PANGAEA, https://doi.org/10.1594/PANGAEA.765164, in supplement to: Lein, AY et al. (2007): Mineralogical and geochemical features of authigenic carbonates on seepings and hydrothermal fields (by the examples of the Black Sea reefs and the mounds of the Lost City field). Translated from Okeanologiya, 2007, 47(4), 577-593, Oceanology, 47(4), 537-553, https://doi.org/10.1134/S000143700704011X, 2007.

Lorrain, A., Pethybridge, H., Cassar, N., Receveur, A., Allain, V., Bodin, N., Bopp, L., Choy, C. A., Duffy, L., Fry, B., Goni, N., Graham, B. S., Hobday, A. J., Logan, J. M., Ménard, F., Menkes, C. E., Olson, R. J., Pagendam, D. E., Point, D., Revill, A. T., Somes, C. J., and Young, J. W.: Trends in tuna carbon isotopes suggest global changes in pelagic phytoplankton communities, Global Change Biology, pp. 458 – 470, https://doi.org/10.1111/gcb.14858, 2020.

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021




MacKenzie, K. M., Robertson, D. R., Adams, J. N., Altieri, A. H., and Turner, B. L.: Carbon and nitrogen stable isotope data from organisms in the Bay of Panama ecosystem, https://doi.org/10.1594/PANGAEA.903842, supplement to: MacKenzie, KM et al. (2019): Structure and nutrient transfer in a tropical pelagic upwelling food web: From isoscapes to the whole ecosystem. Progress in Oceanography, 178, 102145, https://doi.org/10.1016/j.pocean.2019.102145, 2019.

Morée, A. L., Schwinger, J., and Heinze, C.: Southern Ocean controls of the vertical marine $\delta^{13}$C gradient – a modelling study, Biogeosciences, 15, 7205–7223, https://doi.org/10.5194/bg-15-7205-2018, https://doi.org/10.5194/bg-15-7205-2018, 2018.

NOAA's Pacific Marine Environmental Laboratory: Ferret Support, http://ferret.pmel.noaa.gov/Ferret.

Popp, B. N., Takigiku, R., Hayes, J. M., Louda, J. W., and Baker, E. W.: The post-palaeozoic chronology and mechanism of 13C depletion in primary marine organic matter, American Journal of Science, pp. 436–454, 1989.

Popp, B. N., Laws, E. A., Bidigare, R. R., Dore, J. E., Hanson, K. L., and Wakeham, S. G.: Effect of Phytoplankton Cell Geometry on Carbon Isotopic Fractionation, Geochimica et Cosmochimica Acta, 62, 69–77, https://doi.org/10.1016/s0016-7037(97)00333-5, https://doi.org/10.1016/s0016-7037(97)00333-5, 1998.

Rau, G. H., Takahashi, T., and Marais, D. J. D.: Latitudinal variations in plankton $\delta$13C: implications for CO2 and productivity in past oceans, Nature, 341, 516–518, https://doi.org/10.1038/341516a0, https://doi.org/10.1038/341516a0, 1989.

Rau, G. H., Riebesell, U., and Wolf-Gladrow, D.: A model of photosynthetic 13C fractionation by marine phytoplankton based on diffusive molecular CO2 uptake, Marine Ecology Progress Series, 133, 275–285, 1996.

Rocha, C. D. L. and Passow, U.: The Biological Pump, in: Treatise on Geochemistry, pp. 93–122, Elsevier, https://doi.org/10.1016/b978-0-08-095975-7.00604-5, https://doi.org/10.1016/b978-0-08-095975-7.00604-5, 2014.

Rounick, J. S. and Winterbourn, M. J.: Stable carbon isotopes and carbon flow in ecosystems - Measuring $^{13}$C to $^{12}$C ratios can help to trace carbon pathways, https://doi.org/10.2307/1310304, 1986.

Sacket, W. M., Eckelmann, W. R., Bender, M. L., and Bé, A. W. H.: Temperature Dependence of Carbon Isotope Composition in Marine Plankton and Sediments, Science, 148, 235–237, https://doi.org/10.1126/science.148.3667.235, 1965.

Saupe, S. M., Schell, D. M., and Griffiths, W. B.: Carbon-isotope ratio gradients in western arctic zooplankton, Marine Biology, 103, 427–432, https://doi.org/10.1007/BF00399574, 1989.

Schmittner, A. and Somes, C. J.: Complementary constraints from carbon ($^{13}$C) and nitrogen ($^{15}$N) isotopes on the glacial ocean's soft-tissue biological pump, Paleoceanography, pp. 669–693, https://doi.org/10.1002/2015PA002905, 2016.

Silverman, B. W.: Density Estimation for Statistics and Data Analysis, Monographs on Statistics and Applied Probability, Chapman and Hall, 1986.

Suess, E.: Particulate organic carbon flux in the oceans—surface productivity and oxygen utilization, Nature, 288, 260–263, 1980.

Tagliabue, A. and Bopp, L.: Towards understanding global variability in ocean carbon-13, Global Biogeochemical Cycles, 22, n/a–n/a, https://doi.org/10.1029/2007gb003037, https://doi.org/10.1029/2007gb003037, 2008.

Thiede, J., Gerlach, S. A., Altenbach, A., and Henrich, R.: Sedimentation im europaeischen Nordmeer - Organisation und Forschungsprogramm des Sonderforschungsbereiches 313 fuer den Zeitraum 1988-1990, Tech. rep., Kiel University, 1988.

Trull, T. W. and Armand, L. K.: Insights into Southern Ocean carbon export from the $\delta^{13}$C of particles and dissolved inorganic carbon during the SOIREE iron release experiment, Deep Sea Research Part II: Topical Studies in Oceanography, 48, 2655–2680, https://doi.org/10.1016/S0967-0645(01)00013-3, 2001.

Trull, T. W. and Armand, L. K.: $\delta^{13}$C content of particulate organic carbon measured on samples from traps during TANGAROA cruise SOIREE, https://doi.org/10.1594/PANGAEA.807904, 2013a.

https://doi.org/10.5194/essd-2021-159
Preprint. Discussion started: 26 May 2021




Trull, T. W. and Armand, L. K.: $\delta^{13}$C content of fractionated particulate organic carbon measured on samples from traps during TANGAROA cruise SOIREE, https://doi.org/10.1594/PANGAEA.807906, 2013b.

Tuerena, R. E., Ganeshram, R. S., Humphreys, M. P., Browning, T. J., Bouman, H., and Piotrowski, A. P.: Isotopic fractionation of carbon during uptake by phytoplankton across the South Atlantic subtropical convergence, Biogeosciences, 16, 3621–3635, https://doi.org/10.5194/bg-16-3621-2019, 2019.

Verwega, M.-T., Somes, C. J., Tuerena, R. E., and Lorrain, A.: A global marine particulate organic carbon-13 isotope data product, https://doi.org/10.1594/PANGAEA.929931, https://doi.org/10.1594/PANGAEA.929931, 2021.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nature Methods, 17, 261–272, https://doi.org/10.1038/s41592-019-0686-2, 2020.

Voss, M. and von Bodungen, B.: Carbon and nitrogen from mooring NB2, https://doi.org/10.1594/PANGAEA.106805, 2003.

Wada, E., Terazaki, M., Kabaya, Y., and Nemoto, T.: $^{15}$N and $^{13}$C abundances in the Antarctic Ocean with emphasis on the biogeochemical structure of the food web, Deep Sea Research Part A. Oceanographic Research Papers, 34, 829–841, https://doi.org/10.1016/0198-0149(87)90039-2, 1987.

Westerhausen, L. and Sarnthein, M.: $\delta^{13}$C of plankton from surface water (Table A2), https://doi.org/10.1594/PANGAEA.89388, 2003.

Young, J. N., Bruggeman, J., Rickaby, R. E. M., Erez, J., and Conte, M.: Evidence for changes in carbon isotopic fractionation by phytoplankton between 1960 and 2010, Global Biogeochemical Cycles, 27, 505–515, https://doi.org/10.1002/gbc.20045, 2013.