

Review of “MOftern River archivEs of Particulate Organic Carbon: MOREPOC” by Yutian Ke, Damien Calmels, Julien Bouchez, and Cecile Quantin for publication in *Earth System Science Data*.

This dataset is a comprehensive inventory of TOC, $\delta^{13}\text{C}$, F^{14}C , C:N ratios, Al:Si ratios, and important methods related metadata for particulate organic carbon collected and analyzed from rivers around the world. This builds significantly upon previously compiled datasets, which have an order of magnitude fewer datapoints than this new dataset submitted by Ke and colleagues. The publication of this dataset is timely, as the number of studies measuring the geochemistry of fluvial POC has increased over the past decade, and more studies are adopting the dual isotope measurement approach. I am excited about the future studies this dataset will enable.

This dataset seems to be thorough with respect to including all available published data and the dataset includes the relevant parameters and metadata needed to understand how the data were collected. However, there are some issues with the dataset, particularly with respect to reporting measurement uncertainties, naming the variables used to represent the data, and formatting and populating sampling dates. There are also minor issues throughout the text that need to be addressed before this manuscript can be published.

After the dataset and manuscript have been revised to address all issues detailed below, I support publication of this manuscript in *Earth System Science Data*.

Detailed comments:

Dataset:

One significant issue with this dataset is that there are no uncertainties reported for the POC $\delta^{13}\text{C}$, F^{14}C , and radiocarbon age measurements. Analytical uncertainties are typically required when publishing these isotopic measurements, so they should be available in most published datasets. Please add these as columns in the dataset.

The names of the first two columns of the dataset (riv_na and bas_na) are not intuitive. Without looking carefully, I would interpret riv_na to be the name of the sampled river, which bas_na would be the name of the major drainage basin that river is in. To ease use of the dataset, I recommend switching these column headers so that riv_na is the name of the sampled river and bas_na is the name of the larger drainage basin to which the sampled river contributes.

With respect to the river name, I suggest removing the country name from the river name and adding a separate column for the country. In its current format, if someone wants to filter the dataset by specific river, they need to type in the country in parentheses after the river name, which makes data analysis challenging.

CN_mar and as_mar also are not intuitive variable names. CN_ratio and als_i_ratio seem more appropriate and do not exceed the maximum number of characters used for other variables in the dataset.

Rca could also be changed to age_14C.

91 samples do not have dates associated with them, which I assume is because the dates were not reported in the publication. I would appreciate if the authors could confirm whether they double checked the data sources for information on sampling dates.

549 of the reported measurements do not have individual dates, but a range of years (e.g., 1999-2004). This is not very useful for someone who wants to calibrate POC data to points on river hydrographs. It is also unclear whether the reported geochemical measurements are from individual samples, or a calculated composite measurement of multiple samples. This either needs to be explained in the text, or the dataset needs to be edited to breakout the samples into individual dates.

One of the samples (Amazon River at Obidos, 2005) has the latitude and longitude reversed, such that it plots down near Antarctica. It appears that this point was removed from Figure 1, but it is still in the spreadsheet and in the shapefile.

Manuscript text:

Line 7-8: This sentence could be re-organized to digest more easily. I suggest “Riverine transport of particulate organic carbon (POC) associated with terrigenous solids to the ocean has an important role in the global carbon cycle.”

Line 31: “...over geological timescales (>100,000 years).”

Line 45: While SOC in permafrost regions is certainly depleted in ^{14}C because it is old and ^{14}C still decays over time, this organic matter may not have a long turnover time. When frozen, its decomposition rate is zero, but when thawed it may actually have a fast turnover time. We don’t yet have enough data to constrain the turnover time because the turnover time “clock” is only set once the permafrost thaws, which has been occurring more recently.

Line 47: The Carvalhais et al. (2014) study shows total ecosystem carbon turnover times, not necessarily soil organic carbon turnover time and is not constrained by ^{14}C data, so using this reference is a misleading. The Shi et al. (2020) reference is appropriate here.

Line 53: Write “in reservoirs” rather than “at reservoirs.”

Line 66: Suggested edit: “...with 531 reported $\delta^{14}\text{C}$ measurements.”

Line 81: “Decarbonization” is not an appropriate word to use here. I suggest “carbonate removal methods.” It is also unclear what is meant by “acid adopted.” Do you mean type of acid used for removing carbonate? If so, please reword for clarity.

Table 2:

- Decarbonization is not an appropriate term to use here, because it implies removing all carbon. Rather, the authors should use “carbonate removal.”

Line 126: Again, not sure that decarbonization is the right term to use to describe carbonate removal. It implies that all carbon is removed from the samples. It is more straightforward to say “carbonate removal method” or “inorganic carbon removal method.”

Lines 134-136: What are the units associated with time and temperature of acid treatment? This is also not reported in the data spreadsheet.

Line 142: "...consists *of* an extensive dataset..."

Line 146: There should be a comma or colon after 1950, not a quotation mark.

Lines 142-154: Both of these equations have F¹⁴C as an input, but the authors never provide an equation to derive F¹⁴C from an AMS measurements of ¹⁴C/¹²C. I recommend adding the equation for F¹⁴C:

$$F^{14}C = \left[\frac{\left(\frac{^{14}C}{^{12}C} \right)_{sample, -25}}{0.95 \left(\frac{^{14}C}{^{12}C} \right)_{OX1, -19}} \right]$$

Where the denominator is 95% of the ¹⁴C activity of the Oxalic Acid 1 standard material in 1950, and the numerator is corrected for fractionation to a common δ¹³C value of -25 per mil.

Line 170: "range" not "ranges"

Line 259: Are Earth system modelers not a part of the scientific community? This sentence needs to be re-arranged or re-worded. I suggest deleting "as well as Earth system modelers." It would be worth adding a sentence mentioning that having such a dataset to work with can help inform and validate Earth system models, improving our ability to model the global carbon cycle.