Global patterns of radiocarbon depletion in subsoil linked to rock-derived organic carbon

K.E. Grant, R.G. Hilton, V.V. Galy

Supplementary Information

The Supplementary Information includes:

- 1. ISRaD Database Screening and Geological Maps
- 2. Summary of Radiocarbon ($\Delta^{14}$C) Data from ISRaD Soil Profiles
- 3. Measured Soil and River Samples with Known $\text{OC}_{\text{petro}}$ Inputs
- 4. Model Expectations
- 5. Maximum Permissible $\text{OC}_{\text{petro}}$ Input to Soils
- 6. Caveats in the Global Approach
- Supplementary Tables S-1 to S-4
- Supplementary Figures S-1 to S-3
- Supplementary Information References

1. ISRaD Database Screening and Geological Maps

1.1 Database parent material. We used the ISRaD database (ISRaD_data_v1.3.4.2020) of soil radiocarbon measurements (reported as $\Delta^{14}$C) to explore the role of rock organic carbon ($\text{OC}_{\text{petro}}$) inputs. Over 79% of the soils listed in the ISRaD database do not have parent material information or chemistry. Table S-1 shows all listed parent material variables in the ISRaD database, with the number of entries, relative proportion of entries, and final designation we assigned using external geological mapping (see section 1.2). Of the data fields listed in the ISRaD database (a total of 181) we used the ‘lyr_c_org’, ‘lyr_c_tot’, ‘lyr_14c’, ‘lyr_fraction_modern’, ‘pro_name’, ‘pro_parent_material’, ‘pro_treatment’, ‘site_lat’, and ‘site_long’ for this analysis. Other data fields were not used in this study.
Table S-1  Parent Material Categories listed in ISRaD database.

<table>
<thead>
<tr>
<th>ISRaD Parent Material Designation</th>
<th># of entries</th>
<th>Proportion of total (%)</th>
<th>GLiM Designation</th>
</tr>
</thead>
<tbody>
<tr>
<td>'NA'</td>
<td>13,842</td>
<td>79.4</td>
<td>assigned</td>
</tr>
<tr>
<td>'evaporites'</td>
<td>59</td>
<td>0.3</td>
<td>removed</td>
</tr>
<tr>
<td>'igneous extrusive'</td>
<td>163</td>
<td>0.9</td>
<td>IG</td>
</tr>
<tr>
<td>'igneous intrusive'</td>
<td>630</td>
<td>3.6</td>
<td>IG</td>
</tr>
<tr>
<td>'igneous pyroclastic'</td>
<td>229</td>
<td>1.3</td>
<td>IG</td>
</tr>
<tr>
<td>'interbedded'</td>
<td>13</td>
<td>0.1</td>
<td>SED</td>
</tr>
<tr>
<td>'loess'</td>
<td>465</td>
<td>2.7</td>
<td>removed</td>
</tr>
<tr>
<td>'metamorphic'</td>
<td>379</td>
<td>2.2</td>
<td>SED</td>
</tr>
<tr>
<td>'organic'</td>
<td>283</td>
<td>1.6</td>
<td>removed</td>
</tr>
<tr>
<td>'sedimentary-clastic'</td>
<td>1361</td>
<td>7.8</td>
<td>SED</td>
</tr>
<tr>
<td>Total</td>
<td>17,424</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1.2 Assessment of bedrock geology at ISRaD soil profile locations. To assign generalised parent material for each sample in the ISRaD database we used the site location (latitude and longitude) and compared that location to the bedrock geology of the global lithologic map (GLiM) (Hartmann and Moosdorf, 2012). The GLiM is a GIS layer of bedrock or exposed surface rock on the Earth in a 0.5 by 0.5 degree grid space. It is the highest resolution grid for a global surface land model for rock type. The GLiM has 16 simplified lithologic classes of parent material or exposed surface rock. We collated these rock types to form the SED and IG definitions. For the purposes here, we use only these primary lithologic classes, rather than the subclasses, and as such this approach simplifies the heterogeneity that can exist within potential sedimentary rock types.

For samples with parent material field reported with ‘NA’ in the ISRaD database (13,842) we assign a designation of SED or IG based on the co-located bedrock from the GLiM database. For the SED designation, we used five classes (mt: Metamorphic, sc: Carbonate sedimentary rocks, sm: Mixed sedimentary rocks, ss: Siliciclastic sedimentary rocks, su: Unconsolidated sediments) and for the IG we used pa: Acid plutonic rocks, pb: Basic plutonic rocks, pi: Intermediate plutonic rocks, py: pyroclastic, va: Acid volcanic rocks, vi: intermediate volcanic rocks. Excluded from both the sedimentary and igneous datasets were profiles that were designated as on the ic: Ice and glaciers and on wb: waterbodies by the GLiM database. Additionally, any sites that had listed parent material in ISRaD categorised as follows ‘interbedded’, ‘sedimentary-clastic’, and ‘metamorphic’ as SED, while ‘igneous extrusive’, igneous intrusive’, igneous pyroclastic’ are designated as IG. The other parent material designations of ‘evaporates’, ‘loess’, and 'organic' were removed from this analysis. Additionally, in both the SED or IG designations the ‘treatment’ profiles were removed. Hereafter, any soil profile with sedimentary parent material is designated (SED) (397 profiles; 2260 individual entries) and soil profiles formed on igneous parent material are (IG) (160 profiles; 718 entries). We used a total of 557 profiles in the study, with 2978 total horizons.

1.3 Globally modelled mean age and Δ¹⁴C. To assign the modelled mean age and depth integrated Δ¹⁴C at each of our SED or IG profile locations, we use spatially resolved global datasets of soil OC mean age and depth integrated Δ¹⁴C published by Shi et al. (2020). Using locations of each of our previously designated SED and IG profiles, we extract the depth resolved Δ¹⁴C values and modelled mean ages at each of the 397
SED and 160 IG profiles. We find that the distributions of the both the mean age and the depth resolved $\Delta^{14}$C for our designated SED and IG profiles are statistically different using a Kruskal Wallis test for both the subsoils (<30 cm) and the entire averaged profile ($p > 0.05$). This could be explained by a previously unaccounted for input of OC$_{petro}$ in the SED profiles.

2. Summary of Radiocarbon ($\Delta^{14}$C) Data from ISRaD Soil Profiles

2.1 ISRaD radiocarbon summary. For the radiocarbon data, the ISRaD database contains entries reported in either fraction modern (Fm) and $\Delta^{14}$C values. Any entry as Fm was converted to $\Delta^{14}$C using Eq. S-1 and the sampling year reported in the database:

$$\Delta^{14}C = \left[ Fm \times e^{\Delta(1950-Y_c)} - 1 \right] \times 1000 \quad (S-1)$$

where $Y_c$ is the sampling year listed in the database (pro_year). Table S-2 shows the summary statistics for the entire $\Delta^{14}$C database. The $\Delta^{14}$C depth profiles for each soil profile show radiocarbon depletion is strongly linked with depth (Fig. S-2) as has been seen in other studies (Mathieu et al., 2015; He et al., 2016; Shi et al., 2020).

| Table S-2 Radiocarbon signatures ($\Delta^{14}$C) and % OC for SED and IG. |
|---------------------------------|-------|-------|-------|-------|
|                                | SED   | IG    | SED   | IG    |
| **# samples**                  | 2260  | 718   | 2260  | 718   |
| **% of total**                 | 100   | 100   | 100   | 100   |
| **mean**                       | −164.6| −115.2| 7.3   | 8.0   |
| **standard deviation (±s.d.)** | 259.1 | 261.1 | 14.5  | 11.8  |
| **standard error (±s.e.)**     | 5.5   | 9.7   | 0.30  | 0.44  |
| **median**                     | −112.9| −49.9 | 1.0   | 2.5   |
| **minimum**                    | −981.0| −976.5| 0.0   | 0.0   |
| **maximum**                    | 675.2 | 278.8 | 92.8  | 55.0  |
3. Measured Soil and River Samples with Known OCpetro Inputs

We use soil samples and river sediment and suspended particulate samples where OCpetro incorporation was identified using other geochemical measurements. We use a mixture of soil profiles and river samples for two reasons. Firstly, there have been very few studies to look at OCpetro evolution and influence within individual soil profiles. Secondly, we use rivers to understand the catchment wide influence of bedrock influence on the sampling of river OC.

3.1 River inputs of OCpetro. The two river systems, Andean and Himalayan, draining sedimentary rock basins have been extensively studied, so we have a reasonable understanding of OCpetro input to these rivers. For Andean river OC, evolution of N/C and δ13C strongly suggests mixing of OC from soils and bedrocks, which were independently sampled. This mixing of OCpetro and OCbio was confirmed by 14C measurements (Clark et al., 2013, 2017). For Himalayan river OC, Galy et al. (2008) used Raman spectroscopy to identify recycled graphite reaching rivers. This conclusively demonstrated that OCpetro was incorporated into the ‘modern’ river system through weathering. In another study, ramped pyrolysis/oxidation (RPO-14C) coupled to 14C measurements was used to identify a high temperature, 14C ‘dead’ component in the river OC. This technique was used to quantify the rock OC input to the river system (Rosenheim and Galy, 2012).

3.2 Soil inputs of OCpetro. For soil samples, a study done in Taiwan was chosen based on its conclusive examination of OCpetro occurrence during soil paedogenesis, erosion, and weathering. Taiwan soil OC which has been studied extensively using both RPO-14C and lipid biomarkers (abundance, 13C, and 14C isotope composition) to trace OCpetro though a soil weathering profiles, and ultimately record OCpetro oxidation in rapidly eroding soils (Hemingway et al., 2018).
While we also consider black shale weathering profiles as soil, they are more accurately described as an oxidative weathering profile. They are formed by weathering rather than a “top down” soil forming process. These well-studied samples reflect a high [OC] with completely depleted $^{14}$C signature and are a specific case endmember to show $\text{OC}_{\text{petro}}$ can be lost in a profile (Petsch et al., 2001).

4. Model Expectations

4.1. Soil decay model. To explore the range of OC and $\Delta^{14}$C values expected for organic matter undergoing decay in soil, we model loss of OC (lowering [OC], %) as the material ages and the loss of $^{14}$C by radioactive decay (Eq. S-1). Radiocarbon is continually produced in the atmosphere and is incorporated into the biosphere through photosynthesis. Soils become enriched in $^{14}$C via plant inputs. Subsequently, this organic matter can be distributed into multiple soil reservoirs, which can be modelled numerically as individual “pools” (Eq. S-2) in the soil with different residence time. These can reflect different depths in a soil, and/or different physical and/or bio-chemical characteristics which alter the degradation of organic matter:

$$C(t) = \sum_{i=1-3} C_i(t)$$  \hspace{1cm} (S-2)

where $C$ is the total OC remaining through time and is a function of 1 to 3 different pools ($i$) with different residence times. Soil OC decay models with 1, 2 or 3 pools, each with different OC decay constants, $k_i$ were explored. We assume a steady state system and are therefore only interested in the decay or output of the “box”, so we can explicitly solve for the change in the soil decay as:

$$\frac{dC_{\text{Tot}}}{dt} = -k_i \times C_i(t)$$  \hspace{1cm} (S-3)

The metric $1/k_i$ is commonly referred to as the turnover time (Sierra et al., 2017). The model is not trying to fit the data, but instead track the range of expected values of $1/\text{OC}$ and $\Delta^{14}$C. As such, we do not make the model more complex by considering interacting pools (Schuur et al., 2016) and start with $[\text{OC}] = 30\%$ (Fig. S-2), representing plant organic matter inputs, and then follow the trajectory with no additional OC input. These models track the loss of OC over a fixed amount of time. All modelling was done using Matlab version R2019b.

We tested parallel multiple pool models (2 and 3 pools) with turnover times representing “fast”, “slow”, and “inert” OC decay. A “multi-pool” model may better describe a ‘real’ soil system (Gaudinski et al., 2000; Manzoni et al., 2009). In each model, we find that the low [OC] outputs, (i.e. [OC] < 1%) are defined by the pool with the longest turnover time. So, for our study, a 1 pool model was representative of the average decay over long timescales and low total OC outputs, which are most relevant for assessing $\text{OC}_{\text{petro}}$ input in deeper soil. We do not seek to unravel the complexity of rapid decay on shorter time scales. The soil decay model predicts non-linear trajectories in $1/[\text{OC}]$ versus $\Delta^{14}$C, with a steep drop and an asymptotic behaviour for a given turnover time (Fig. 1b).

4.2 Endmember mixing model. To explore the $\Delta^{14}$C and [OC] values produced by mixing $\text{OC}_{\text{petro}}$ and $\text{OC}_{\text{bio}}$, we use binary mixing models to calculate a range of $\text{OC}_{\text{petro}}$ inputs. A two-endmember mixing model describing these two endmembers is defined as follows:
\[
\Delta^{14}C_{\text{mix}} = f_{\text{petro}} \times \Delta^{14}C_{\text{petro}} + f_{\text{bio}} \times \Delta^{14}C_{\text{bio}}
\]  

(S-4)

\[
1 = f_{\text{petro}} + f_{\text{bio}}
\]  

(S-5)

\[
\% \text{OC}_{\text{mix}} = \% \text{OC}_{\text{petro}} + \% \text{OC}_{\text{bio}}
\]  

(S-6)

\[
1 = \frac{\% \text{OC}_{\text{petro}}}{\% \text{OC}_{\text{mix}}} + \frac{\% \text{OC}_{\text{bio}}}{\% \text{OC}_{\text{mix}}}
\]  

(S-7)

\[
\Delta^{14}C_{\text{mix}} = \frac{\% \text{OC}_{\text{petro}}}{\% \text{OC}_{\text{mix}}} \times \Delta^{14}C_{\text{petro}} + \frac{\% \text{OC}_{\text{mix}} - \% \text{OC}_{\text{petro}}}{\% \text{OC}_{\text{mix}}} \times \Delta^{14}C_{\text{bio}}
\]  

(S-8)

where Equation S-4 states that the bulk \(\Delta^{14}C\) value of the sampled soil (\(\Delta^{14}C_{\text{mix}}\)) is a function of the fraction of OC_{petro} (\(f_{\text{petro}}\)) which by definition has \(\Delta^{14}C_{\text{petro}} = -1000 \text{‰}\) and the fraction of OC_{bio} (\(f_{\text{bio}}\)) with an \(\Delta^{14}C_{\text{bio}}\) close to the atmospheric value. This OC_{bio} value is variable and based on the ‘modern’ input of OC to the soil, in this case \(\Delta^{14}C_{\text{bio}}\) is typically \(\geq 100 \text{‰}\). In this instance all the organic carbon in the single soil sample is either petrogenic or biospheric (Eq. S-5), so the fraction of each must equal 1. This can be rewritten to reflect the known % OC measurements where the total sample % OC value is % OC_{mix} and the either measured or model % OC_{petro} and % OC_{bio} contributing to the % OC_{mix} (Eq. S-6). We can then substitute our Equation S-7 into Equation S-4 to solve for OC_{petro} (Eq. S-8) because we know the measured OC_{mix} and the bulk \(\Delta^{14}C_{\text{mix}}\) and \(\Delta^{14}C_{\text{petro}}\) and can estimate \(\Delta^{14}C_{\text{bio}}\) from atmospheric values.

Here we consider the OC_{bio} endmember is \(\Delta^{14}C = +200 \text{‰}\) and the OC_{petro} endmember is, by definition, \(\Delta^{14}C = -1000 \text{‰}\). An aged OC_{bio} component is not considered here for simplicity, but if included it would bound the ‘grey’ region with a steeper slope and less area. The distinction in mixing of OC_{bio} with a 200 %‰ and 0 %‰ is only important at relatively young or enriched \(^{14}C\) values. We consider OC_{petro} concentrations from 1 % to 0.1 %, and \(^{14}C\) depletion below \(\Delta^{14}C < -400 \text{‰}\). This defines a mixing zone (“grey region”) where OC_{petro} mixing can generate data in the 1/[OC] and \(\Delta^{14}C\) space.
**Figure S-2**  (a) Binary mixing model (lines) between OC\textsubscript{bio} ($\Delta^{14}C = +200 \%o$) and OC\textsubscript{petro} (varying from 1 \% to 0.1 \%, $\Delta^{14}C = -1000 \%o$). The grey shaded area is where >50 \% is OC\textsubscript{petro}.  B. (b) The predicted evolution of $1/[OC]$ and radiocarbon activity ($\Delta^{14}C$) for single pool OC decay models with turnover times (years) shown.

### 5. Maximum Permissible OC\textsubscript{petro} Input to Soils

**5.1 OC\textsubscript{petro} in each horizon.** Using the binary mixing model analysis described above (Eqs. S-2 to S-8) we calculate the maximum OC\textsubscript{petro} input from rocks into each individual horizon using carbon concentration (\% OC) and measured radiocarbon signature ($\Delta^{14}C$) of SED samples from the ISRaD database. We do this for samples that have radiocarbon values less than 0 \%o because we assume values near modern would have little influence by OC\textsubscript{petro} inputs. We solve Equation S-4 for the fraction of OC\textsubscript{petro} that could be contributing to decrease the sample $\Delta^{14}C$, to the measured $\Delta^{14}C\textsubscript{mix}$. We use a biospheric OC (OC\textsubscript{bio}) $\Delta^{14}C$ value is determined by the average atmospheric $\Delta^{14}C$ value at the year reported in the ISRaD database. The atmospheric value is taken from Levin and Kromer (2004). Inclusion of an aged OC\textsubscript{bio} endmember would lower the OC\textsubscript{petro} contribution. We calculate a $f_{\text{petro}}$ and OC\textsubscript{petro} for each $\Delta^{14}C$ value <0 \%o in our SED dataset (Table S-3). To provide a first constraint on this, we assume that OC\textsubscript{bio} aging dominates the surface of SED soil profiles where $\Delta^{14}C > -200 \%o$. For samples with $\Delta^{14}C < -200 \%o$ ($n = 877$), the average \% OC = 2.3 ± 0.3 and $\Delta^{14}C = -430 ± 6.4 \%o$. If the OC\textsubscript{petro} fraction was 0.38, \textit{i.e.} ~38 \% of the total OC (Fig. 3B), the residual OC\textsubscript{bio} would have $\Delta^{14}C = -80.6 \%o$. 
Table S-3  Analysis of maximum % OC<sub>petro</sub> input to SED samples.

<table>
<thead>
<tr>
<th></th>
<th>% OC&lt;sub&gt;petro&lt;/sub&gt;: Δ&lt;sup&gt;14&lt;/sup&gt;C ≤ 0 ‰</th>
<th>f&lt;sub&gt;petro&lt;/sub&gt;: Δ&lt;sup&gt;14&lt;/sup&gt;C ≤ 0 ‰</th>
</tr>
</thead>
<tbody>
<tr>
<td># samples</td>
<td>1565</td>
<td>1565</td>
</tr>
<tr>
<td>mean</td>
<td>0.85</td>
<td>0.38</td>
</tr>
<tr>
<td>standard deviation (±s.d.)</td>
<td>2.29</td>
<td>0.21</td>
</tr>
<tr>
<td>standard error (±s.e.)</td>
<td>0.06</td>
<td>0.01</td>
</tr>
<tr>
<td>median</td>
<td>0.19</td>
<td>0.36</td>
</tr>
<tr>
<td>minimum</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>maximum</td>
<td>22.5</td>
<td>1.0</td>
</tr>
</tbody>
</table>

5.2 Comparison to global sedimentary rocks. We compare maximum % OC<sub>petro</sub> contribution calculated in the SED soil horizons to the known global distribution of OC in sedimentary rocks that has been previously published in Partin et al. (2013). We expect our calculated % OC<sub>petro</sub> to be an overestimation in our soils, but compared to % OC in global sedimentary rocks, where by definition all OC is OC<sub>petro</sub>, our average maximum calculated % OC<sub>petro</sub> for soils is 6.5 times less than the average % OC in globally distributed sedimentary rocks. This is evidence for widespread oxidation of OC<sub>petro</sub> during soil formation and weathering.

Table S-4  Global Sedimentary Rock % OC values taken from Partin et al. (2013).

<table>
<thead>
<tr>
<th></th>
<th>Total Dataset (% OC)</th>
<th>Zeros removed (% OC)</th>
<th>&lt;20 % OC</th>
</tr>
</thead>
<tbody>
<tr>
<td># of samples</td>
<td>2730</td>
<td>2544</td>
<td>2451</td>
</tr>
<tr>
<td>mean</td>
<td>5.13</td>
<td>5.50</td>
<td>4.77</td>
</tr>
<tr>
<td>standard deviation (±s.d.)</td>
<td>6.34</td>
<td>6.41</td>
<td>5.18</td>
</tr>
<tr>
<td>standard error (±s.e.)</td>
<td>0.12</td>
<td>0.13</td>
<td>0.10</td>
</tr>
<tr>
<td>median</td>
<td>2.4</td>
<td>2.8</td>
<td>2.6</td>
</tr>
<tr>
<td>minimum</td>
<td>0</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>maximum</td>
<td>39.3</td>
<td>39.3</td>
<td>19.83</td>
</tr>
</tbody>
</table>
6. Caveats in the Global Approach

6.1 Preprocessing of soil samples: inorganic carbon removal. Our study requires the inclusion of samples with both measured total carbon and measured radiocarbon values. Total carbon was used in place of organic carbon values if organic carbon values were not reported. Most samples only reported total carbon measurements, without any specific information on the presence of inorganic carbon. In typical soil/sediment radiocarbon preparation, inorganic carbon (carbonate species) are removed prior to graphitisation via acid leaching or fumigation (Bao et al., 2019). We assume that these published radiocarbon measurements would have undergone these pre-procedures if it was suspected detrital carbonates may be present, and thus that the dataset is not impacted by inorganic carbon inputs.

6.2 Preprocessing of soil samples: sieving. The standard protocol for soil profile sample collection includes the removal of rock fragments, either using a sieve or by hand (Soil Survey Staff, 2014). In this way, the ISRaD database may contain lower rock-derived inputs than is actually present in the environment. Nevertheless, sand, silt and clay materials in soils can still contain OCpetro (Agnelli et al., 2002; Hemingway et al., 2018; Kalks et al., 2021). Rock fragments could be an important source of OC at all depths in a soil profile (Trumbore and Zheng, 1996; Agnelli et al., 2002). It is traditional in soil studies to remove >2 mm fragments of the soil during sample processing before any geochemical analysis is completed. The >2 mm is reported by mass, but usually not measured otherwise. This practice may skew the radiocarbon age if those rock fragments have a measurable portion of OC derived from the parent material especially if the parent material contains OCpetro. The argument is that these particles are too large to play an active geochemical role, however this portion of the soil is actively weathering to create the <2 mm fraction. Most of the ISRaD database is likely to be comprised of measurement of the <2 mm portion of the soil, which would mean our assessment of OCpetro inputs is conservative in the SED profiles.
6.3 Presence of loess. We filtered the ISRaD database to remove profiles which mention Loess as the parent material (Table S-1). However, some of the N/A parent material classifications may be on loess, as this material is common in the USA Midwest and in East Asia. We have chosen not to actively filter further, recognising that loess material could contain rock OC, and it is a sedimentary structure, albeit not directly linked to the bedrock. The GLiM clearly states that it is measuring the parent material, likely exposed at the surface, but not the soil. As stated before, we have not used the subcategories for the GLiM due to simplicity. There is a chance that we would have a mixture of sources as the parent material as dust inputs, erosive material, or bedrock may be different than the exact surface exposure (Dere et al., 2013). Therefore, in the very deep soils on “unconsolidated sediments” as is determined by the GLiM, as well as their secondary classification layer (lo: Loess), so while these sites may not be directly related to the “bedrock” the parent material in certainly sedimentary, and likely could contain ‘rock’ OC that has been weathered and transported.

6.4 Presence of clay minerals. The statistical differences in Δ14C values for SED and IG samples support a role of OCpetro input. However, there could be other reasons that SED samples are more 14C-depleted, related to the presence of clay minerals. Mineral protection of OC can increase the relative proportion of 14C-depleted OC in a sample. In volcanic, igneous basaltic soils, OC can be protected >10,000 years due to the short-range order mineral “gel” formed between co-precipitated OC and mineral matrix (Torn et al., 1997). While this might suggest igneous parent material has a larger capacity to protect OC, thus resulting in more 14C depleted OC, it is well established that more traditional clay minerals hosted in sedimentary rocks such as illite and kaolinite have a large capacity to sorb OC as well. These clay minerals can be formed as weathering products in granites, but are also common in sedimentary rocks, such as shales. Shales are formed through marine or lacustrine detrital sediment with high OC values during deposition. These clay minerals (both in size and mineralogy) can contribute to the stable OC content in shales, and this would suggest that shales can protect all types of OC through increased clay content (Basu et al., 2019). It is reasonable that shale could protect even young carbon for extended periods of time. We cannot rule out a role of clay mineralogy and its difference between SED and IG profiles, and this represents an important topic for future work, the simpler explanation is that OCpetro inputs at SED sites explain the [OC] and Δ14C distributions.

6.5 Implications of OCpetro input for SOC storage. OCpetro input could potentially alter SOC storage mechanisms in two ways. Generally, soil OC persistence mechanisms are either physical/chemical protection or intrinsic reactivity (Lehmann and Kleber, 2015). OCpetro is inherently part of the mineral fraction of the soil as it is a part of the bedrock. It likely will be both physically and chemically bonded to the mineral material. We hypothesise that the majority of the OCpetro would be in the heavy or mineral associated fraction, depending on the method used to fractionate or separate the soil pools (Lützow et al., 2006). OCpetro in the mineral associated fraction would drive the measured radiocarbon age to be older while the actual age would be younger, just as we found in the bulk soils. In addition, OCpetro would be subject to oxidation as the surface area of the rock material is exposed to oxygen and water. Another hypothesis is this OCpetro could possibly contribute to additional sites for OCbio accumulation due to organic-organic bonds (Possinger et al., 2020) and contribute to additional OCbio storage. It is very difficult to further comment on these processes without direct studies of the mineral bound soil carbon pool, and these remain avenues for future research.
Supplementary Information References


