nature geoscience

Article

https://doi.org/10.1038/s41561-024-01476-4

Global patterns of organic carbon transfer and accumulation across the land–ocean continuum constrained by radiocarbon data

Received: 20 October 2023

Accepted: 29 May 2024

Published online: 21 June 2024

Check for updates

Chenglong Wang[1,5] ✉, Yifei Qiu[1,5], Zhe Hao[2], Junjie Wang[3], Chuchu Zhang[1], Jack J. Middelburg[3], Yaping Wang[1] & Xinqing Zou[1,4] ✉

Radiocarbon ($\Delta^{14}C$) serves as an effective tracer for identifying the origin and cycling of carbon in aquatic ecosystems. Global patterns of organic carbon (OC) $\Delta^{14}C$ values in riverine particles and coastal sediments are essential for understanding the contemporary carbon cycle, but are poorly constrained due to under-sampling. This hinders our understanding of OC transfer and accumulation across the land–ocean continuum worldwide. Here, using machine learning approaches and >3,800 observations, we construct a high-spatial resolution global atlas of $\Delta^{14}C$ values in river–ocean continuums and show that $\Delta^{14}C$ values of river particles and corresponding coastal sediments can be similar or different. Specifically, four characteristic OC transfer and accumulation modes are recognized: the old–young mode for systems with low river and high coastal sediment $\Delta^{14}C$ values; the young–old and old–old modes for coastal systems with old OC accumulation receiving riverine particles with high and low $\Delta^{14}C$ values, respectively; and the young–young mode with young OC for both riverine and coastal deposited particles. Distinguishing these modes and their spatial patterns is critical to furthering our understanding of the global carbon system. Specifically, among coastal areas with high OC contents worldwide, old–old systems are largely neutral to slightly negative to contemporary atmospheric carbon dioxide (CO_2) removal, whereas young–old and old–young systems represent CO_2 sources and sinks, respectively. These spatial patterns of OC content and isotope composition constrain the local potential for blue carbon solutions.

Coastal margins play a key role in the global carbon (C) cycle, constituting <10% of the entire ocean area, but contributing >90% to the overall organic carbon (OC) burial in global oceans[1,2]. Coastal margins serve as the interface between terrestrial and oceanic C pools and receive a diverse mixture of organic compounds from both allochthonous terrestrial inputs and autochthonous marine primary production[3,4]. However, OC with different origins exhibits distinct signatures, reactivities and ages[4,5], therefore responding differently to remobilization and alteration processes in coastal oceans[5,6]. For instance, rivers transport ~200 megatonnes of particulate OC per year to oceans[7], of which 55–80% is remineralized along continental margins[3]. Furthermore, anthropogenic activities have substantially altered regional and global C cycling (for example, production, transport, preservation and burial of OC) along river–ocean continuums[8,9]. Therefore, obtaining accurate information about the nature and burial of OC in global coastal sediments is challenging but crucial for developing robust C budgets and predicting future C dynamics.

[1]School of Geography and Ocean Science, Ministry of Education Key Laboratory for Coast and Island Development, Nanjing University, Nanjing, China. [2]School of Marine Science and Engineering, Nanjing Normal University, Nanjing, China. [3]Department of Earth Sciences, Utrecht University, Utrecht, the Netherlands. [4]Collaborative Innovation Center of South China Sea Studies, Nanjing University, Nanjing, China. [5]These authors contributed equally: Chenglong Wang, Yifei Qiu. ✉e-mail: clwang@nju.edu.cn; zouxq@nju.edu.cn

https://doi.org/10.1038/s41561-024-01476-4

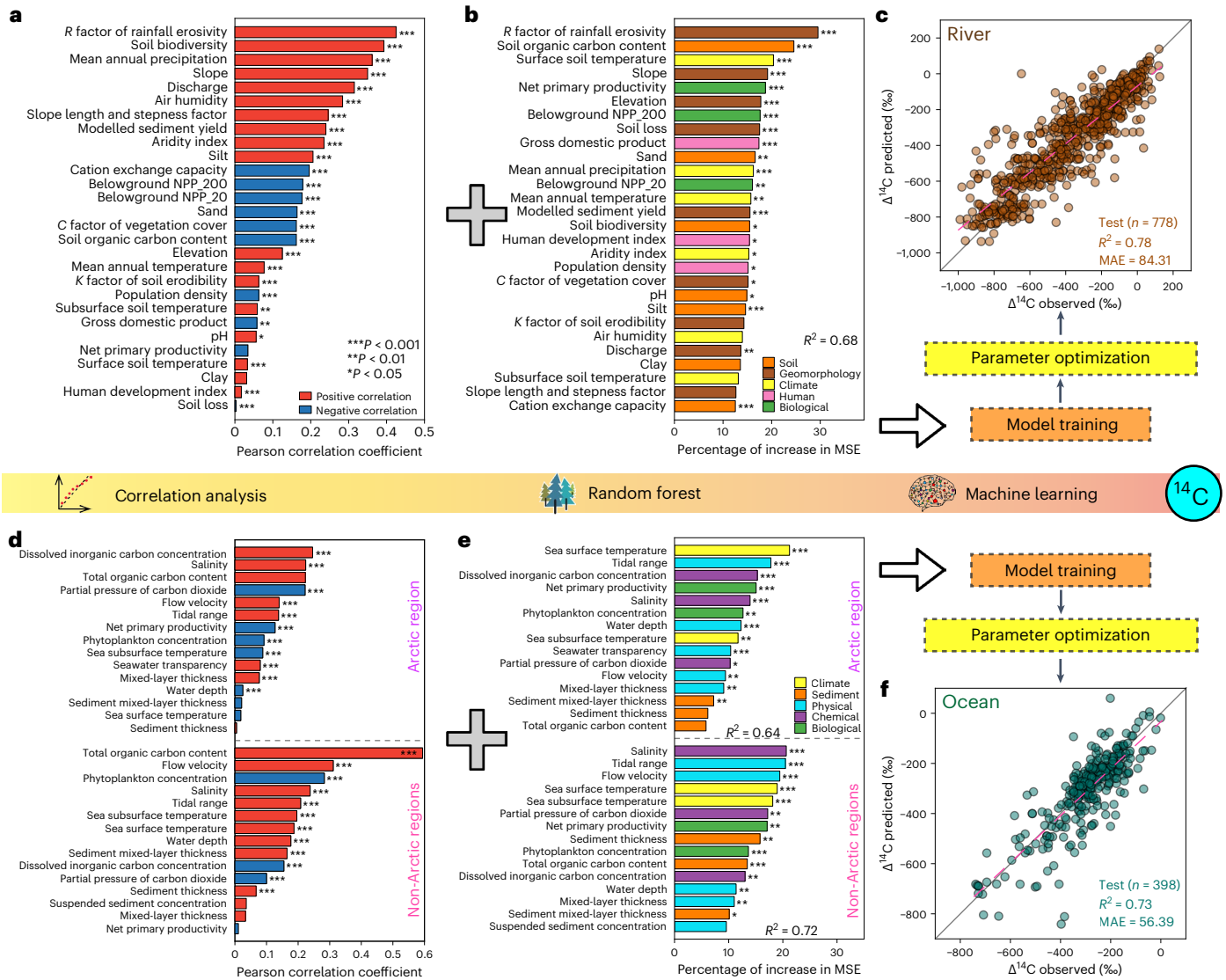


Fig. 1 | Environmental variables driving $\Delta^{14}C$ values of riverine particles and coastal sediments. a–f, Importance of environmental variables in driving $\Delta^{14}C$ values of riverine particles (**a**–**c**) and coastal sediments (**d**–**f**) based on correlation analysis (**a** and **d**), the random forest approach (**b** and **e**) and plots of overall prediction performance using the feature-selected SVM for $\Delta^{14}C$ values (**c** and **f**). $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$. The percentage of increase in the mean square error (MSE) represents the importance of environmental variables in driving $\Delta^{14}C$ using the random forest approach. Belowground NPP_20, belowground net primary productivity at 20 cm depth; Belowground NPP_200, belowground net primary productivity at 200 cm depth; R^2, coefficient of determination; MAE, mean absolute error.

Radiocarbon ($\Delta^{14}C$) of OC has emerged as a powerful tool for investigating contemporary aquatic C biogeochemistry[10–12], offering insights into the average age of the local OC mixed with various organic compounds of different ages[4]. More-negative $\Delta^{14}C$ values (that is, ^{14}C-depleted) indicate older ^{14}C ages (that is, a predominance of old OC, possibly mixed with some modern OC) and vice versa. The $\Delta^{14}C$ value not only serves as a tracer of OC sources with diverse ages (for example, ^{14}C-enriched young photosynthetic C, ^{14}C-depleted aged petrogenic C and pre-aged C somewhere in between)[4,12], but also provides a window on the dynamics of C transfer between surface C reservoirs on Earth[13]. Recent studies have elucidated basin-scale variations of $\Delta^{14}C$ in some riverine and marine C pools (for example, refs. 14,15) and attempts have been made to compile available $\Delta^{14}C$ information to obtain their global patterns[13,16–21].

However, although the number of $\Delta^{14}C$ measurements has increased greatly, available data cover only a fraction of global river basins and coastal zones and are unevenly distributed spatially, with frequent measurements at some locations and none in others (Extended Data Fig. 1). So far, a high-resolution dataset of $\Delta^{14}C$ with global coverage is lacking for both river particles and coastal sediments. Moreover, available observations in river basins and coastal sediments are usually disconnected, complicating a systematic analysis of $\Delta^{14}C$ along river–ocean continuums. Scaling up unevenly distributed site-level $\Delta^{14}C$ observations to the global scale at a high spatial resolution for both global rivers and coastal zones is imperative to better constrain the global C cycle in a spatially explicit manner.

Machine learning has emerged lately as a powerful tool for scaling up site-level observations to high-resolution global patterns, particularly in the fields of Earth system science[22], terrestrial C and nitrogen biogeochemistry[23,24] and marine sediment geochemistry[25,26]. These machine learning applications offer a new approach to the challenging high-resolution mapping of $\Delta^{14}C$ values in global river–ocean continuums.

In this Article, we outline the compilation of $\Delta^{14}C$ data for 2,559 observations (737 locations) of riverine particles and 1,325 observations (1,325 locations) of coastal surface sediments (depths <5 cm) worldwide

https://doi.org/10.1038/s41561-024-01476-4

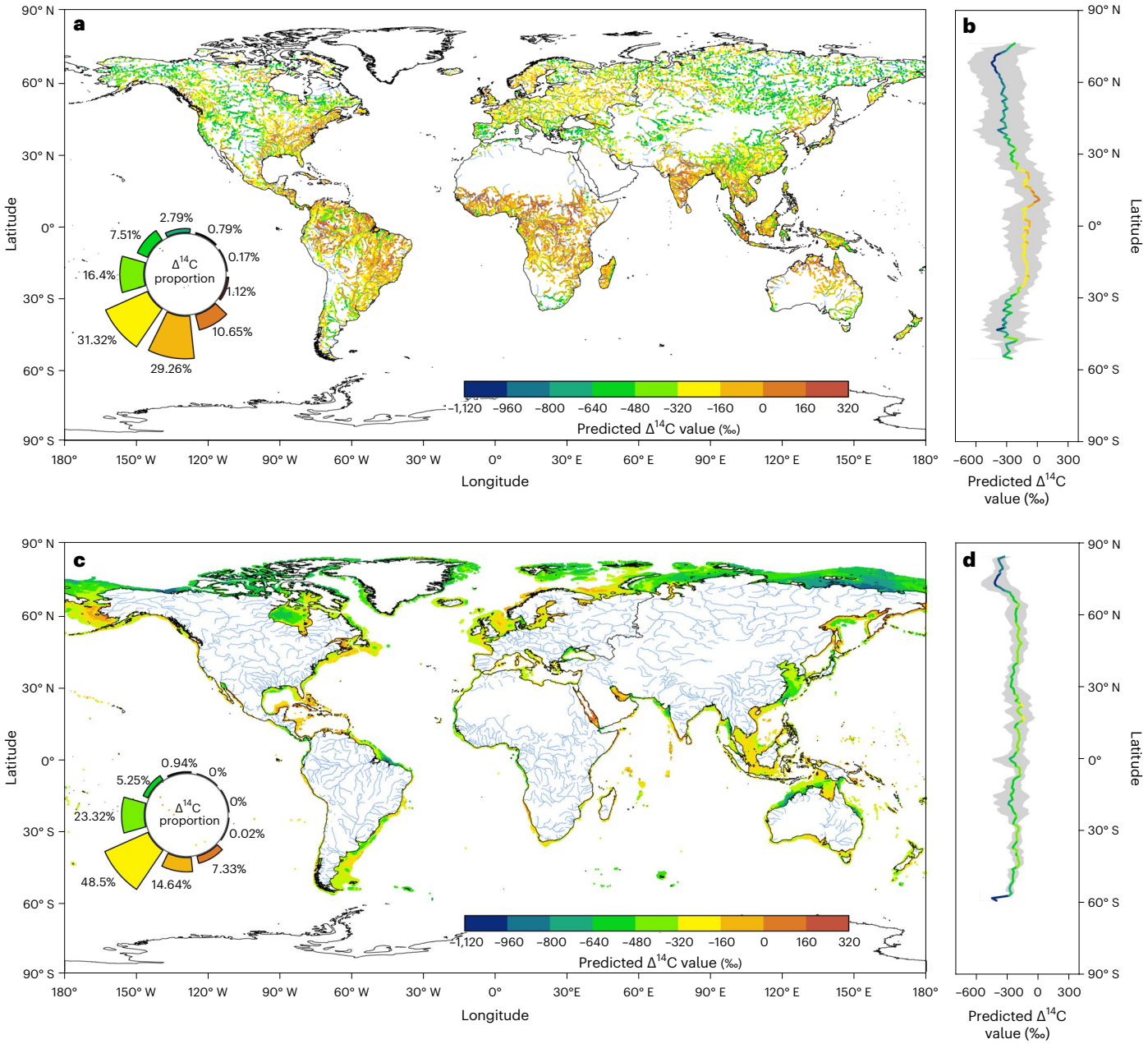


Fig. 2 | Global distributions and latitudinal patterns of predicted $\Delta^{14}C$ values in riverine particles and coastal sediments. a–**d**, Global distributions (**a** and **c**) and latitudinal patterns (**b** and **d**) of predicted $\Delta^{14}C$ values in riverine particles (**a** and **b**) and coastal sediments (**c** and **d**). The map in **a** is shown at river orders 1–7, as defined by the classical ordering system of HydroBASINS (https://www.hydrosheds.org/products/hydrobasins) and includes 22,442 predicted values (Methods). The map in **c** is shown at a spatial resolution of 10′ × 10′ and includes 99,807 predicted values (Methods). The coloured lines in **b** and **d** represent mean values and the grey shading represents s.d.

(Extended Data Fig. 1) from published literature and databases, including Circum-Arctic Sediment Carbon Database (CASCADE), Modern River Archives of Particulate Organic Carbon (MOREPOC) and Modern Ocean Sediment Archive and Inventory of Carbon (MOSAIC)[19–21] (Methods). We also compiled extensive data for environmental variables known to regulate C delivery and accumulation (Supplementary Table 1). These data were used to train and test machine learning models (Fig. 1) and to generate a high-resolution global map of $\Delta^{14}C$ values in river–ocean continuums. We also performed machine learning simulations for total OC (TOC) contents and $\delta^{13}C$ values of global coastal sediments. Using a combination of the high-resolution global patterns of TOC contents and $\Delta^{14}C$ and $\delta^{13}C$ values, we identified accumulation hotspots for modern and aged OC from marine and terrestrial sources in coastal oceans worldwide.

Global mapping of OC $\Delta^{14}C$ values in river–ocean continuums

Using the machine learning approach and compiled data, we generated a high-resolution $\Delta^{14}C$ atlas for river particles and coastal sediments with good agreement with observations (Methods and Supplementary Discussion). This atlas provides high-spatial resolution $\Delta^{14}C$ values with complete global coverage (Fig. 2). We further converted $\Delta^{14}C$ values into ^{14}C ages to obtain high-resolution global patterns of ^{14}C ages (Extended Data Fig. 2 and Methods).

Spatially, the predicted $\Delta^{14}C$ values of global riverine particles exhibit variability and vary sharply with latitude (Fig. 2a,b). Positive or less-negative $\Delta^{14}C$ values are predominantly located in South America, Africa and Southern and Southeastern Asia (Fig. 2a). The most-negative $\Delta^{14}C$ values are mainly found in the Arctic permafrost

https://doi.org/10.1038/s41561-024-01476-4

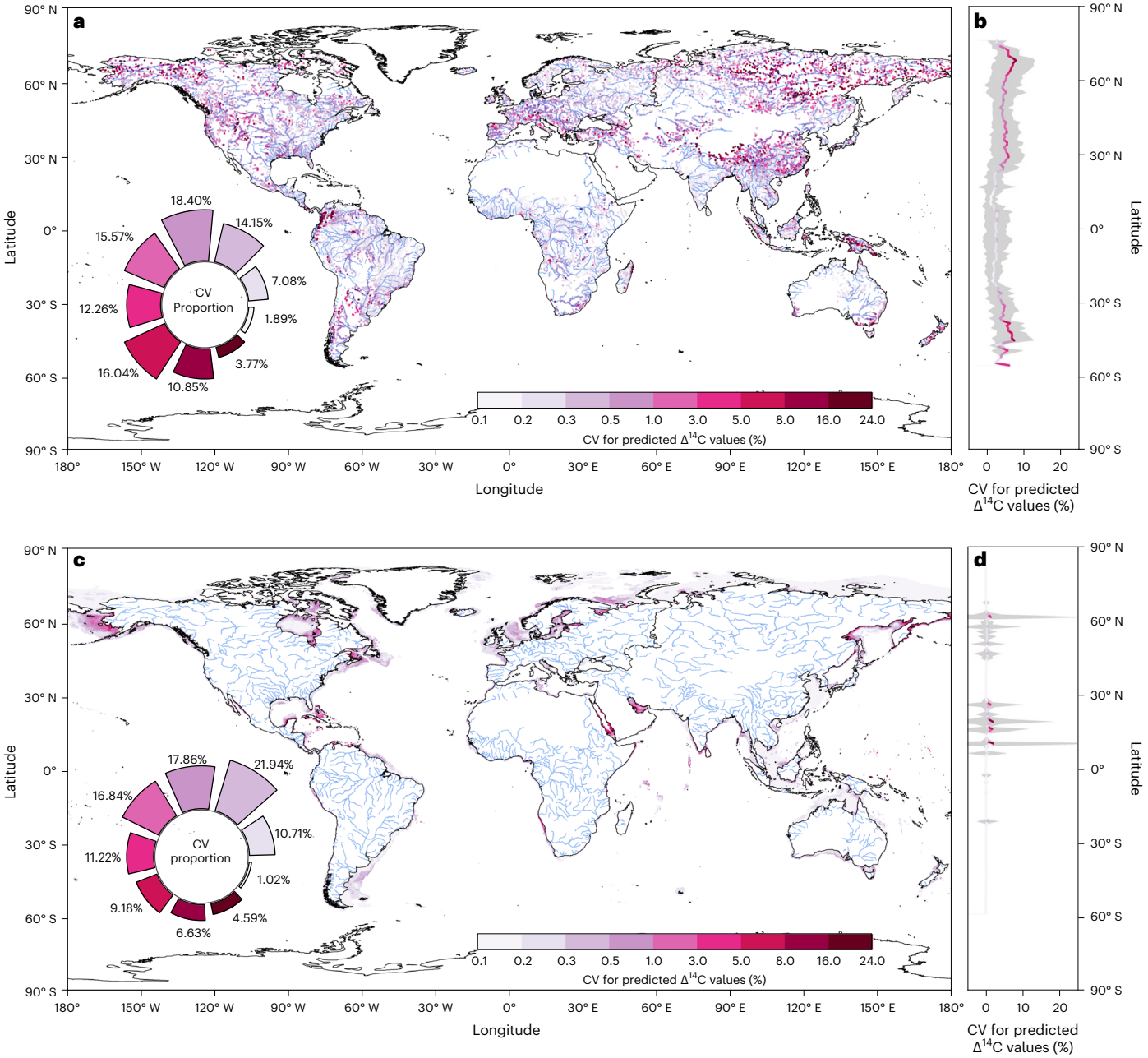


Fig. 3 | Global distributions and latitudinal patterns of predicted uncertainty of $\Delta^{14}C$ values in riverine particles and coastal sediments. a–**d**, Global distributions (**a** and **c**) and latitudinal patterns (**b** and **d**) of predicted uncertainty in $\Delta^{14}C$ values in riverine particles (**a** and **b**) and coastal sediments (**c** and **d**). The map in **a** is shown at river orders 1–7, as defined by the classical ordering system of HydroBASINS (https://www.hydrosheds.org/products/hydrobasins) and includes 22,442 predicted values (Methods). The map in **c** is shown at a spatial resolution of 10′ × 10′ and includes 99,807 predicted values (Methods). The coefficient of variation (CV) is used to represent the uncertainty of machine learning models. The coloured lines in **b** and **d** represent mean values and the grey shading represents s.d.

region, highlands such as the Qinghai–Tibet and Mongolian Plateaus and mountainous regions such as Taiwan Island and the Western United States (Fig. 2a). Along the latitudinal gradient, $\Delta^{14}C$ values are more negative in high-latitude regions (beyond 60° N or 60° S) and less negative (or more positive) in low-latitude regions (between 30° N and 30° S) (Fig. 2b).

This spatial variability also exists in the predicted $\Delta^{14}C$ values of global coastal sediments (Fig. 2c,d). Positive or less-negative $\Delta^{14}C$ values are primarily in the Subarctic shelf, Sunda shelf, Caribbean Sea and parts of the African coast (Fig. 2c). The most-negative $\Delta^{14}C$ values are mainly in wide shelf areas such as the Arctic shelf, East China Sea and northern shelf of Australia, as well as near mouths of rivers such as the Amazon, Huanghe, Indus, Mississippi and Irrawaddy estuaries (Fig. 2c). The latitudinal trend of $\Delta^{14}C$ of global coastal sediments is similar to that of global riverine particles, with more-negative values in high-latitude regions and less-negative values in low-latitude regions (Fig. 2d).

The predictive uncertainty of $\Delta^{14}C$ values in global riverine particles ranges from 0.01–22.74%, with an average of 3.15 ± 2.27% (Fig. 3a). Similarly, the predictive uncertainty of $\Delta^{14}C$ values in coastal sediments ranges from 0.13–15.34%, with an average of 2.04 ± 2.70% (Fig. 3c). One important reason for the predictive uncertainties is the low spatial density of sampling, together with the high spatial variability of observed $\Delta^{14}C$ values. High predictive uncertainties occur in regions with limited or no observations (for example, high-altitude or high-latitude regions,

https://doi.org/10.1038/s41561-024-01476-4

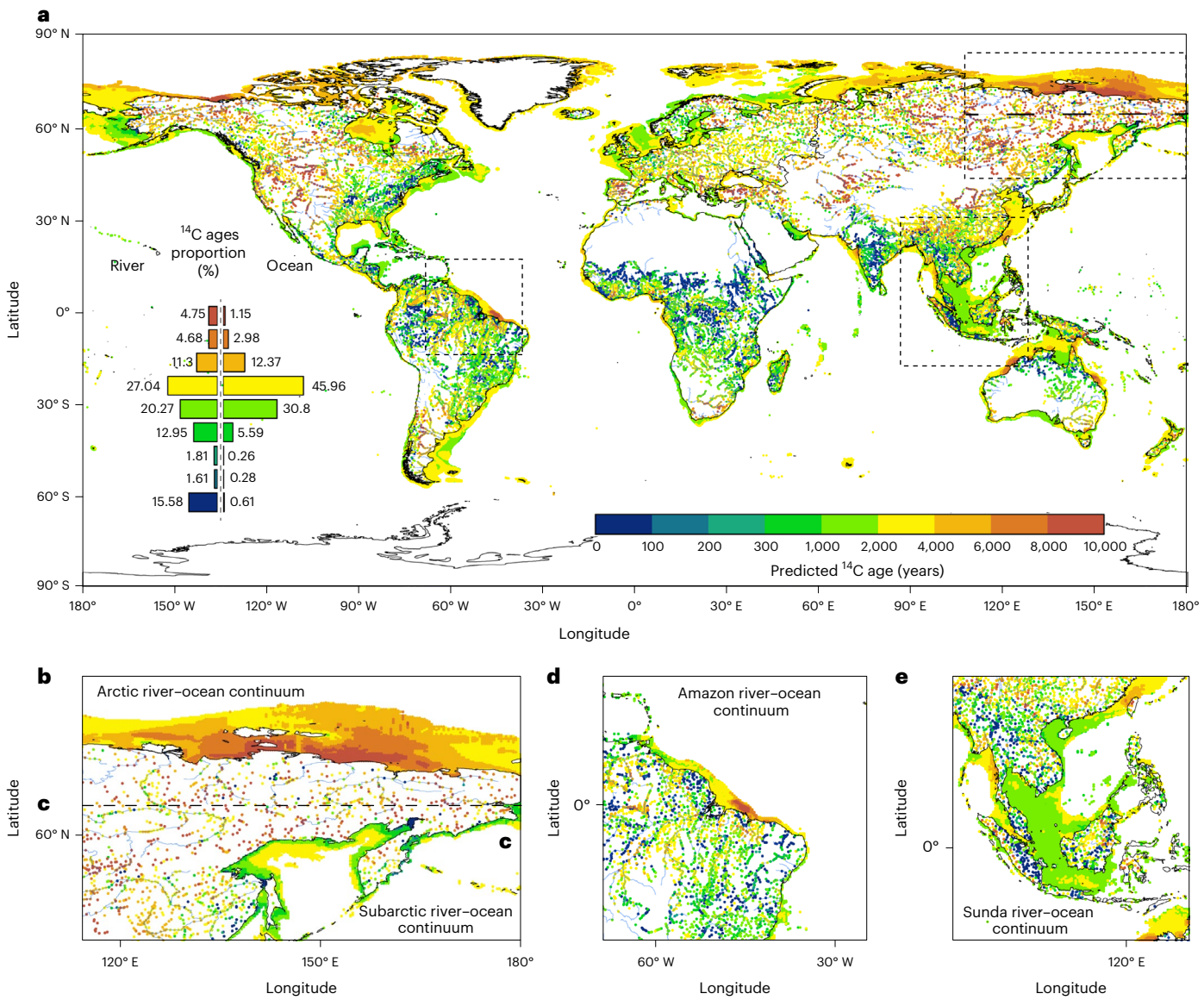


Fig. 4 | Global distribution of ^{14}C ages in riverine particles and coastal sediments worldwide and for a range of typical sub-regions. a–e, Global distribution of ^{14}C ages in both riverine particles and coastal sediments worldwide (**a**) and magnified for a range of typical sub-regions: the Arctic (**b**), Subarctic (**c**), Amazon river–ocean continuum (**d**) and Sunda river–ocean continuum (**e**). Panels **b**–**d** are magnified from the sections outlined by dashed boxes in **a**. The global map in **a** shows river orders 1–7, as defined by the the classical ordering system of HydroBASINS (https://www.hydrosheds.org/products/hydrobasins) and includes 22,442 predicted riverine particle values (Methods). For coastal sediments, 99,807 predicted values are included in **a** at a spatial resolution of 10′ × 10′ (Methods). The predicted ^{14}C ages were converted from the predicted $\Delta^{14}C$ values in Fig. 2 following the approach described in the Methods.

such as the Arctic continent and Subarctic coastal ocean; Fig. 3a,c and Extended Data Fig. 1a,b). In addition, $\Delta^{14}C$ values may vary with river discharge[13,17], while small, mountainous rivers in the mid-latitudes (for example, Taiwan Island and Southeast Asia) are under-sampled at high-discharge times (Fig. 3a and Extended Data Fig. 1). Along the latitudinal gradient, the predictive $\Delta^{14}C$ uncertainties for riverine particles show variability, with higher values in high-latitude regions and much lower values in low-latitude regions (Fig. 3b). For coastal sediments, the latitudinal variation in uncertainties is characterized by higher values in the Northern Hemisphere and lower values in the Southern Hemisphere (Fig. 3d).

Overall, the utilization of machine learning techniques, together with available observations, has enabled high-resolution prediction of $\Delta^{14}C$ values in global river–ocean continuums, with uncertainty, accuracy and spatial coverage that compare favourably with existing databases[19–21]. The global patterns of $\Delta^{14}C$ values in riverine particles and coastal sediments produced in this study can be applied to investigation of the C cycle regionally and globally.

Critical modes of river–coastal OC ^{14}C ages worldwide

Four distinct modes emerged in this study (Fig. 4): the old–young mode describes old ^{14}C ages in riverine particles accompanied by young ^{14}C ages in corresponding coastal sediments; the young–old mode describes young ^{14}C ages in riverine particles coupled with old ^{14}C ages in corresponding coastal sediments; the young–young mode describes young ^{14}C ages in both riverine particles and corresponding coastal sediments; and the old–old mode describes old ^{14}C ages in both riverine particles and corresponding coastal sediments.

The old–young mode is common in the Subarctic river–ocean continuum (Fig. 4c), western coast of the United States and coast of South Africa (Fig. 4a and Table 1). The annual export of riverine

https://doi.org/10.1038/s41561-024-01476-4

Table 1 | Five modes of OC ^{14}C ages in river–ocean continuums worldwide, their typical regions and their role in the contemporary coastal C cycle

^{14}C age mode	Typical regions	Coastal OC accumulation hotspots	Contribution to removing contemporary atmospheric C
Old–young	Southeastern coast of Australia, western coast of United States and coast of South Africa	Yes, for young OC accumulation	Positive
Young–old	Amazon, Changjiang, Indus, Irrawaddy, Mississippi and Fly estuaries, and northwestern and northeastern Australia	Yes, for old OC accumulation	Negative
Young–young	Caribbean Sea, Sunda river–ocean continuum, western coast of Mexico and South China Sea	Yes/no, for young OC accumulation	Positive/negative/no effect—vulnerable to change
Old–old A	Arctic shelf and Argentine continental shelf	Yes, for old OC accumulation	Negative
Old–old B	Mountainous river–ocean continuums of Taiwan and eastern Australia	Bypass, non-hotspots	No to (slightly) negative

particles of old ^{14}C ages (mainly petrogenic C) to these coastal regions is limited[3], with pulsed transport and burial of riverine OC occurring over days to weeks, associated with storms but possibly not captured in the sampling dates[3]. The young (that is, recently produced) OC in these coastal sediments may mainly originate from marine primary productivity[27] (Supplementary Fig. 1). The hotspots of young ^{14}C ages on the western coast of the United States (for example, offshore California), coast of South Africa (for example, coast of Cape Peninsula) and Subarctic regions (for example, Shelikhov Gulf and Sakhalinskiy Bay) match the locations of reported upwelling zones[28–30], where upwelling of nutrient-rich water stimulates in situ primary productivity[27] (Supplementary Figs. 1 and 2). However, in the Subarctic river–ocean continuum the predicted ^{14}C ages are highly uncertain (Fig. 3c) due to the scarcity of local observations.

The young–old mode is primarily in low-latitude regions, such as the Amazon, Congo and Fly river–ocean continuums (Fig. 4d). These systems have extensive floodplains with large standing stocks of biomass, which can be transported by rivers to coastal zones during flooding periods[31]. This pattern is also in some subtropical to temperate river–ocean continuums, such as the Changjiang, Indus, Irrawaddy and Mississippi and their estuaries (Fig. 4a and Table 1). Rivers in these regions typically receive OC from terrestrial ecosystems and have large in situ algal production; high temperature and high precipitation enhance terrestrial and freshwater primary production (Supplementary Fig. 3). Eglinton et al.[18] proposed that the ages of riverine biospheric C are positively correlated with the turnover time and ^{14}C ages of soil OC. Negative relationships have been observed globally between the logarithmic OC turnover time and temperature or precipitation[32,33]. Thus, soil OC in low-latitude regions (with high temperature and high precipitation) tends to have shorter turnover times and younger ^{14}C ages[31,34]. In their corresponding coastal zones, marine primary productivity may also be high due to large riverine nutrient inputs (Supplementary Fig. 1). However, ^{14}C ages in these coastal sediments are old and do not reflect the young ^{14}C ages in river particles or the high marine primary production. This is because the young terrestrial, freshwater and marine OC is preferentially degraded in the highly dynamic shallow environments with extensive sediment reworking, as is reflected by their thick sediment mixed layers (SMLs)[26]; for example, a maximum SML thickness of ~200 cm in the Amazon Estuary and large SML thicknesses of ~30 cm in other aforementioned coastal regions (Supplementary Fig. 4). The thick SMLs are primarily attributed to physical perturbation and bioturbation, which cause continuous resuspension–redeposition loops.

The young–young mode primarily occurs in low-latitude regions, typically in the Sunda river–ocean continuum (Fig. 4e), Caribbean coastal regions, western coast of Mexico and South China Sea (Fig. 4a and Table 1). As for the young–old mode, rivers in these regions receive large amounts of OC from terrestrial primary production and have high freshwater production, leading to young ^{14}C ages for river particles. These rivers also transport substantial nutrients from land to sea, enhancing marine primary production (Supplementary Figs. 1 and 5). The ^{14}C ages in their corresponding coastal sediments are also young, which differs from those in the young–old mode. This may be due to the lower hydrodynamics in these coastal zones, resulting in rather stable sedimentary environments and thin SMLs (Supplementary Fig. 4). Therefore, younger OC originating from both autochthonous (for example, phytoplankton detritus) and allochthonous sources (for example, plant debris and phytoplankton detritus) can be effectively deposited and preserved. Moreover, such stable hydrodynamics triggers frequent hypoxic events through oxygen consumption during OC decay and limitations in water–sediment exchange[35]. The hypoxic conditions may further limit bioturbation and microbial respiration and contribute to the preservation of young OC[35,36], as can be seen in the western coast of Mexico, Caribbean coastal regions and Northwest European Shelf (Fig. 4a and Supplementary Fig. 6).

The old–old mode can be divided into two sub-modes: the old–old A mode, which is common in the Arctic river–ocean continuum (Fig. 4b) with a wide shelf; and the old–old B mode in mountainous river–ocean continuums in Taiwan Island and eastern Australia with a narrow shelf (Fig. 4a and Table 1). In the old–old A mode, despite high uncertainties in the predicted ^{14}C ages in Arctic rivers due to limited data coverage (Extended Data Fig. 1), the ^{14}C ages of river particles exported by Arctic rivers are overall older than those in mid- to low latitudes. Such old riverine ^{14}C ages are primarily attributed to the ^{14}C-depleted soils, such as permafrost and Yedoma in Arctic river basins[34]. For example, a recent study[37] showed that the majority of terrestrial OC in the circum-Arctic region originates from near-surface soils (61%) and permafrost (30%). Although another recent study suggests that the warming Arctic may enhance the export of riverine aquatic biomass production[38], it is important to recognize that most of this biomass produced in aquatic environments may be degraded during cross-shelf transport[39]. The marine primary production in Arctic coastal regions tends to be low because of low temperature, low nutrient inputs and low water transparency (Supplementary Figs. 3, 5 and 7). In the old–old B mode, the old ^{14}C ages in coastal sediments are attributed to the substantial input of aged petrogenic OC transported by mountainous river draining areas of high erodibility[40,41].

Implications for coastal OC accumulation and present CO_2

Previous studies have explored the global distribution of sedimentary OC in coastal margins[25,42,43], primarily focusing on OC content. Information regarding OC source, composition and burial potential remains very limited. Traditionally, OC accumulation is discussed in terms of the balance between OC input/production on the one hand and OC degradation/export on the other. OC accumulating in coastal systems is then considered a sink of carbon dioxide (CO_2). However, such a mass balance approach largely ignores the source and ageing

of OC, and this might bias the inferred implications for contemporary atmospheric CO_2 levels[14].

To further elucidate coastal C dynamics, the TOC contents and $\delta^{13}C$ values of global coastal sediments were predicted using similar machine learning techniques (Methods), with satisfying performance of model training and testing (Supplementary Table 2). Using a combination of our high-resolution data of global coverage for TOC contents, $\delta^{13}C$ values and ^{14}C ages, we identified the hotspots of OC accumulation potential in coastal oceans worldwide by accounting for not only OC quantity, but also OC source, composition and ambient environmental conditions. The hotspots of high coastal OC content reported by Bianchi et al.[44] were also found in our study (Extended Data Fig. 3a). Moreover, we identified three different types of hotspots of coastal OC accumulation in surface sediments (hereafter referred to as OC accumulation).

Coastal regions with both high TOC content and the old–young riverine–coastal ^{14}C age mode are the primary hotspots for young OC accumulation (Table 1 and Extended Data Fig. 3a). These regions exhibit less-negative $\delta^{13}C$ values (Extended Data Fig. 3c), indicating that their OC primarily originates from autochthonous marine primary production. This OC origin is also supported by the spatial overlap of OC accumulation hotspots with upwelling zones (supporting high primary production; Supplementary Figs. 1 and 2) and oxygen minimum zones (supporting efficient young OC preservation; Supplementary Fig. 6)[36]. Accumulation of young OC in these hotspots has a positive effect on the removal of contemporary atmospheric CO_2. This result is also consistent with the sea–air CO_2 flux synthesized by Roobaert et al.[45], in which the southeastern coast of Australia and coast of South Africa are a sink of atmospheric CO_2 (Supplementary Fig. 8).

Coastal oceans with high TOC content and the young–old riverine–coastal ^{14}C age mode are primary hotspots for old OC (pre-aged OC and aged petrogenic OC) accumulation (Table 1 and Extended Data Fig. 3a). The $\delta^{13}C$ values of OC in these coastal sediments are predominantly more negative (Extended Data Fig. 3c), indicating that the OC buried here primarily originates from terrestrial inputs. This is supported by the results of *n*-alkanes in the Changjiang Estuary[46], as well as the more frequently detected petroleum (aged petrogenic OC) compared with biogenic hydrocarbons by mooring observations in northeastern Australia[47]. The highly dynamic conditions due to physical disturbances (reflected by thick SMLs) induce degradation of labile OC (both young riverine material and locally produced marine OC) in these coastal regions[3,5], as is reflected by their coastal CO_2 effluxes to the atmosphere (particularly the Amazon Estuary and northeastern Australia; Supplementary Fig. 8). Consequently, the coastal accumulation of old OC does not contribute to contemporary atmospheric CO_2, but the degradation of young OC causes CO_2 emissions, collectively forming a source of contemporary atmospheric CO_2.

Organic-rich sediments in the old–old A mode are another hotspot of old OC accumulation. In the Arctic region, the OC in coastal sediments is mainly from the erosion of permafrost[37], along with minor inputs from aquatic biomass that would probably degrade during cross-shelf transport[39]. This is re-affirmed by the more-negative $\delta^{13}C$ values in these coastal sediments (Extended Data Fig. 3c) and biomarker evidence on the predominant terrestrial source and in situ marine OC degradation[48]. Moreover, the local marine primary production is low (Supplementary Fig. 1) because of low temperature, limited nutrient inputs and low water transparency (Supplementary Figs. 3, 5 and 7). However, not all terrestrial-sourced old OC along these river–coastal continuums is preserved. Molecular degradation proxies indicate that ~1.7 Gg yr^{-1} of old OC in the Arctic shelf is degraded during cross-shelf transport[49], making systems of this type negatively impact the removal of contemporary atmospheric CO_2.

Furthermore, we also identify regions characterized by relatively low TOC content and the young–young riverine–coastal ^{14}C age mode, whose roles in the contemporary C cycle are variable (Table 1 and Extended Data Fig. 3a). The relatively less-negative $\delta^{13}C$ values of OC in most of these regions indicate a dominant source from marine primary production, except for the nearshore South China Sea with relatively more-negative $\delta^{13}C$ values (Extended Data Fig. 3c). This is supported by the combination of terrestrial- and marine-sourced OC in the South China Sea shown by multi-proxy molecular biomarker analyses[50]. The OC in riverine particles in these tropical regions is mainly composed of plant and phytoplankton debris[18,51], thus showing young ^{14}C ages (Fig. 4). This riverine input of young OC, together with young OC from marine primary production, contributes to the young coastal OC in the relatively stable sedimentary environment. These young–young systems can represent a sink or source for atmospheric CO_2; for instance, the sea-air CO_2 flux density atlas shows a CO_2 sink of the northern shelf of the South China Sea (preservation of terrestrial-/riverine-produced OC) and a CO_2 source of the Sunda shelf (consumption of terrestrial OC; Supplementary Fig. 8). Irrespective of whether the C is imported or locally produced, the role of these regions as the source or sink of atmospheric CO_2 depends on the OC balance of the local system, which is vulnerable to global warming because of the labile nature of the young OC. In warmer waters, the temperature-dependent metabolic rates of heterotrophic bacteria increase, thereby accelerating remineralization of the young OC. This process has been used to explain the elevated organic matter recycling efficiency and decreased OC burial in warm climates[52].

In contrast, coastal margins characterized by low TOC content and the old–old B riverine–coastal ^{14}C age mode (Table 1) accumulate only a small amount of C and play a minor role in contemporary CO_2. For instance, OC in the mountainous river–ocean continuums in Taiwan Island and eastern Australia mainly originate from weathering and erosion of bedrocks[40,53]. In the coastal zone of eastern Australia, the in situ marine OC production is very low (Supplementary Fig. 1). In the Taiwan Strait, the relatively enriched ^{13}C and old ^{14}C ages, together with low TOC contents (Extended Data Fig. 3a,b), indicate a dominant contribution of aged petrogenic OC from Taiwan Island[40], rather than marine OC, which rapidly degrades in the highly dynamic environment during transport (Supplementary Figs. 1 and 2). Degradation of old (and sometimes young) OC lowers the TOC contents in these regions, with a slightly negative impact on removing contemporary atmospheric CO_2.

Our study provides new insights into the spatial patterns of global coastal OC accumulation potential by combining machine learning approaches with comprehensive observational data for $\Delta^{14}C$, $\delta^{13}C$ and TOC and their environmental drivers. The high-resolution global maps of $\Delta^{14}C$ values and OC fate in river–ocean continuums from this study, if incorporated, can substantially improve the robustness of C cycling prediction and climate change projections in Earth system models and have far-reaching implications for developing effective zero-CO_2 strategies and national C budgets, including blue C stocks. The characteristic patterns of ^{14}C ages of riverine particles and corresponding coastal sediments demonstrate the relative importance of terrestrial inputs and marine primary production on coastal OC and how ocean margin C budgets relate to factors such as temperature, precipitation and sedimentary dynamics. Our results also point out critical regions with poor data availability (for example, high-altitude or high-latitude regions), necessitating further investigation efforts to understand their local OC dynamics and improve their prediction accuracy. Notably, this study only focuses on particulate OC, while dissolved OC (DOC) also accounts for a large fraction of C in river–-ocean continuums[2]. Including DOC dynamics in future modelling efforts (for example, machine learning or process-based modelling) will further enhance our understanding of global C cycling, in addition to our findings on C dynamics of riverine particles and coastal sediments.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information,

acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41561-024-01476-4.

References

1. Berner, R. A. Burial of organic carbon and pyrite sulfur in the modern ocean; its geochemical and environmental significance. *Am. J. Sci* **282**, 451–473 (1982).
2. Hedges, J. I. & Keil, R. G. Sedimentary organic-matter preservation—an assessment and speculative synthesis. *Mar. Chem.* **49**, 81–115 (1995).
3. Blair, N. & Aller, R. C. The fate of terrestrial organic carbon in the marine environment. *Annu. Rev. Mar. Sci.* **4**, 401–423 (2012).
4. Eglinton, T. I. Variability in radiocarbon ages of individual organic compounds from marine sediments. *Science* **277**, 796–799 (1997).
5. Canuel, E. A. & Hardison, A. K. Sources, ages, and alteration of organic matter in estuaries. *Annu. Rev. Mar. Sci.* **8**, 409–434 (2016).
6. Bianchi, T. The role of terrestrially derived organic carbon in the coastal ocean: a changing paradigm and the priming effect. *Proc. Natl Acad. Sci. USA* **108**, 19473–19481 (2011).
7. Beusen, A. H. W., Dekkers, A. L. M., Bouwman, A. F., Ludwig, W. & Harrison, J. Estimation of global river transport of sediments and associated particulate C, N, and P.*Glob. Biogeochem. Cycles* **19**, GB4S05 (2005).
8. Bauer, J. E. et al. The changing carbon cycle of the coastal ocean. *Nature* **504**, 61–70 (2013).
9. Regnier, P., Resplandy, L., Najjar, R. G. & Ciais, P. The land-to-ocean loops of the global carbon cycle. *Nature* **603**, 401–410 (2022).
10. Druffel, E. R. M. & Bauer, J. E. Radiocarbon distributions in Southern Ocean dissolved and particulate organic matter. *Geophys. Res. Lett.* **27**, 1495–1498 (2000).
11. Raymond, P. A. & Bauer, J. E. Use of ^{14}C and ^{13}C natural abundances for evaluating riverine, estuarine, and coastal DOC and POC sources and cycling: a review and synthesis. *Org. Geochem.* **32**, 469–485 (2001).
12. Galy, V. et al. Efficient organic carbon burial in the Bengal fan sustained by the Himalayan erosional system. *Nature* **450**, 407–410 (2007).
13. Galy, V., Peucker-Ehrenbrink, B. & Eglinton, T. I. Global carbon export from the terrestrial biosphere controlled by erosion. *Nature* **521**, 204–207 (2015).
14. Galy, V., Beyssac, O., France-Lanord, C. & Eglinton, T. I. Recycling of graphite during Himalayan erosion: a geological stabilization of carbon in the crust. *Science* **322**, 943–945 (2008).
15. Bao, R. et al. Widespread dispersal and aging of organic carbon in shallow marginal seas. *Geology* **44**, 791–794 (2016).
16. Griffith, D. R., Martin, W. & Eglinton, T. I. The radiocarbon age of organic carbon in marine surface sediments. *Geochim. Cosmochim. Acta* **74**, 6788–6800 (2010).
17. Marwick, T. R. et al. The age of river-transported carbon: a global perspective. *Glob. Biogeochem. Cycles* **29**, 122–137 (2015).
18. Eglinton, T. I., Galy, V., Hemingway, J., Feng, X. J. & Bao, H. Y. Climate control on terrestrial biospheric carbon turnover. *Proc. Natl Acad. Sci. USA* **118**, e2011585118 (2021).
19. Ke, Y. T., Calmels, D., Bouchez, J. & Quantin, C. MOdern River archivEs of Particulate Organic Carbon: MOREPOC. *Earth Syst. Sci. Data* **14**, 4743–4755 (2022).
20. Martens, J. et al. CASCADE—The Circum-Arctic Sediment CArbon DatabasE. *Earth Syst. Sci. Data* **13**, 2561–2572 (2021).
21. Van der Voort, T. et al. Mosaic (modern ocean sediment archive and inventory of carbon): a (radio)carbon-centric database for seafloor surficial sediments. *Earth Syst. Sci. Data* **13**, 2135–2146 (2021).
22. Reichstein, M. et al. Deep learning and process understanding for data-driven Earth system science. *Nature* **566**, 195–204 (2019).
23. Tao, F. et al. Microbial carbon use efficiency promotes global soil carbon storage. *Nature* **618**, 981–985 (2023).
24. Feng, M. et al. Overestimated nitrogen loss from denitrification for natural terrestrial ecosystems in CMIP6 Earth System Models. *Nat. Commun.* **14**, 3065 (2023).
25. Lee, T. R., Wood, W. T. & Phrampus, B. J. A machine learning (kNN) approach to predicting global seafloor total organic carbon. *Glob. Biogeochem. Cycles* **33**, 37–46 (2019).
26. Song, S. et al. A global assessment of the mixed layer in coastal sediments and implications for carbon storage. *Nat. Commun.* **13**, 4903 (2022).
27. Dai, Y. et al. Coastal phytoplankton blooms expand and intensify in the 21st century. *Nature* **614**, 280–284 (2023).
28. Kämpf, J. On preconditioning of coastal upwelling in the eastern Great Australian Bight. *J. Geophys. Res. Atmos.* **115**, C12071 (2010).
29. Huang, Z. & Wang, X. H. Mapping the spatial and temporal variability of the upwelling systems of the Australian south-eastern coast using 14-year of MODIS data. *Remote Sens. Environ.* **227**, 90–109 (2019).
30. Liao, F. L., Liang, X. F., Li, Y. & Spall, M. A. Hidden upwelling systems associated with major western boundary currents. *J. Geophys. Res. Oceans* **127**, e2021JC017649 (2022).
31. Hoek, W. J. V. et al. Exploring spatially explicit changes in carbon budgets of global river basins during the 20th century. *Environ. Sci. Technol.* **55**, 16757–16769 (2021).
32. Carvalhais, N. et al. Global covariation of carbon turnover times with climate in terrestrial ecosystems. *Nature* **514**, 213–217 (2014).
33. Fan, N. X. Global apparent temperature sensitivity of terrestrial carbon turnover modulated by hydrometeorological factors. *Nat. Geosci.* **15**, 985–994 (2022).
34. Shi, Z. et al. The age distribution of global soil carbon inferred from radiocarbon measurements. *Nat. Geosci.* **13**, 555–559 (2020).
35. Breitburg, D. et al. Declining oxygen in the global ocean and coastal waters. *Science* **359**, eaam7240 (2018).
36. Middelburg, J. J. Reviews and syntheses: to the bottom of carbon processing at the seafloor. *Biogeosciences* **15**, 413–427 (2018).
37. Martens, J., Wild, B., Semiletov, I. P., Dudarev, O. V. & Gustafsson, Ö. Circum-Arctic release of terrestrial carbon varies between regions and sources. *Nat. Commun.* **13**, 5858 (2023).
38. Behnke, M. I. et al. Aquatic biomass is a major source to particulate organic matter export in large Arctic rivers. *Proc. Natl Acad. Sci. USA* **120**, e2209883120 (2023).
39. Bröder, L., Tesi, T., Andersson, A., Semiletov, I. P. & Gustafsson, Ö. Bounding cross-shelf transport time and degradation in Siberian–Arctic land–ocean carbon transfer. *Nat. Commun.* **9**, 806 (2018).
40. Hilton, R. G., Galy, A., Hovius, N., Horng, M.-J. & Chen, H. The isotopic composition of particulate organic carbon in mountain rivers of Taiwan. *Geochim. Cosmochim. Acta* **74**, 3164–3181 (2010).
41. Blair, N. et al. The persistence of memory: the fate of ancient sedimentary organic carbon in a modern sedimentary system. *Geochim. Cosmochim. Acta* **67**, 63–73 (2003).
42. Seiter, K., Hensen, C., Schrötter, J. & Zabel, M. Organic carbÿon content in surface sediments—defining regional provinces. *Deep Sea Res. I* **51**, 2001–2026 (2004).
43. Atwood, T. B., Witt, A., Mayorga, J. S., Hammill, E. & Sala, E. Global patterns in marine sediment carbon stocks. *Front. Mar. Sci.* **7**, 165 (2020).
44. Bianchi, T. S. et al. Centers of organic carbon burial and oxidation at the land–ocean interface. *Org. Geochem.* **115**, 138–155 (2018).
45. Roobaert, A., Regnier, P., Landschützer, P. & Laruelle, G. G. A novel sea surface pCO_2-product for the global coastal ocean resolving trends over 1982–2020. *Earth Syst. Sci. Data* **16**, 421–441 (2024).

46. Wang, C. L. et al. Anthropogenic perturbations to the fate of terrestrial organic matter in a river-dominated marginal sea. *Geochim. Cosmochim. Acta* **333**, 242–262 (2022).
47. Burns, K., Volkman, J. K., Cavanagh, J. E. & Brinkman, D. Lipids as biomarkers for carbon cycling on the northwest shelf of Australia: results from a sediment trap study. *Mar. Chem.* **80**, 103–128 (2003).
48. Bröder, L. et al. Historical records of organic matter supply and degradation status in the East Siberian Sea. *Org. Geochem.* **91**, 16–30 (2016).
49. Bröder, L., Andersson, A., Tesi, T., Semiletov, I. P. & Gustafsson, Ö. Quantifying degradative loss of terrigenous organic carbon in surface sediments across the Laptev and East Siberian Sea. *Glob. Biogeochem. Cycles* **33**, 85–99 (2019).
50. Udoh, E. C. et al. Distribution characteristics of terrestrial and marine lipid biomarkers in surface sediment and their implication for the provenance and palaeoceanographic application in the northern South China Sea. *Mar. Geol.* **452**, 106899 (2022).
51. Mayorga, E. et al. Young organic matter as a source of carbon dioxide outgassing from Amazonian rivers. *Nature* **436**, 538–541 (2005).
52. Li, Z. Y., Zhang, Y. G., Torres, M. & Mills, B. Neogene burial of organic carbon in the global ocean. *Nature* **613**, 90–95 (2023).
53. Hilton, R. G. & West, A. J. Mountains, erosion and the carbon cycle. *Nat. Rev. Earth Env.* **1**, 284–299 (2020).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2024

Methods

Data source and processing

Most of the OC $\Delta^{14}C$ values of riverine particles worldwide were collected from the MOREPOC database[19]. The compiled $\Delta^{14}C$ values from the database were counterchecked with the reported values in the original references to ensure data accuracy. We also searched keywords, including '$\Delta^{14}C$ values/radiocarbon/OC ^{14}C content' and '$\Delta^{14}C$ values/radiocarbon/OC ^{14}C content of riverine particles' in the Web of Science, ResearchGate and Google Scholar and added to our database the $\Delta^{14}C$ data from the most recent publications (that is, those not in MOREPOC). Finally, data of 2,559 observations from 737 sampling sites distributed globally were compiled (Extended Data Fig. 1a).

The $\Delta^{14}C$ values of coastal sediments were mainly collected from the CASCADE and MOSAIC 1.0 published databases[20,21]. Similarly, we added the latest literature-reported data by searching using the Web of Science, ResearchGate and Google Scholar for keywords such as '$\Delta^{14}C$ values/radiocarbon/OC ^{14}C content' and '$\Delta^{14}C$ values/radiocarbon/OC ^{14}C content of sediments'. In the end, we compiled a global dataset of 1,325 coastal sites with $\Delta^{14}C$ values (Extended Data Fig. 1b). In addition, 4,496 sites of TOC contents and 1,554 sites of $\delta^{13}C$ values in global coastal sediments were collected from MOSAIC 1.0 for further analysis.

Data of all environmental variables were collected from published databases (Supplementary Table 1), with 28 variables used for riverine OC $\Delta^{14}C$ value projections and 15 for marine OC $\Delta^{14}C$ value projections. Pearson correlation analysis showed a weak to moderate correlation between pairs of variables within the full set of variables (Extended Data Figs. 4 and 5). Climatic variables that control primary productivity, soil microbial communities and surface weathering and erosion are potential drivers for $\Delta^{14}C$ values of riverine particles[18,32,54], including the aridity index, air humidity, mean annual precipitation, mean annual temperature, surface soil temperature and subsurface soil temperature. Geomorphology is considered an important factor that controls weathering and erosion and influences river export of particulate OC with various $\Delta^{14}C$ values[13,53]. To characterize the effect of geomorphology, variables of elevation, slope, modelled sediment yield, discharge, soil loss, *R* factor of rainfall erosivity, *K* factor of soil erodibility, *C* factor of vegetation cover, slope length and steepness factor were included. Soil properties may influence the degradation and preservation of OC by affecting microbial activities and may change the $\Delta^{14}C$ values of soil OC. The variables used to describe soil properties include clay, silt, sand, soil OC content, pH, cation exchange capacity and soil biodiversity. Anthropogenic activities have been demonstrated to drastically perturb the terrestrial C pool and further influence the export of soil OC[18,55]. To represent various anthropogenic perturbations, variables of population density, human development index and gross domestic product were used. Primary productivity indexes such as net primary productivity and belowground net primary productivity (at 20 and 200 cm depths) are also considered important factors of the $\Delta^{14}C$ values of riverine particles and were used in this study.

The environmental variables used to predict the $\Delta^{14}C$ values of coastal sediments were grouped into physical properties, chemical properties, climate properties, primary productivity indexes and sedimentary properties (Supplementary Table 1). The physical properties included flow velocity, tidal range, mixed-layer thickness, suspended sediment concentration and water depth. Among them, flow velocity, tidal range, mixed-layer thickness and water depth are hydrodynamic parameters that potentially influence OC degradation or preservation in coastal zones by influencing the transport and exchange of OC and oxygen, as well as their oxygen exposure time[39,56]. For instance, a high flow velocity, large tide range and thick mixed layer may prolong the oxygen exposure time and hence the interaction between oxygen and OC, which accelerates the degradation of OC. Chemical properties such as salinity, p_{CO_2} and dissolved inorganic C may influence photosynthesis and microbial activities[57–59] and further impact the $\Delta^{14}C$ values of coastal sediments. Climate properties such as sea surface temperature and sea subsurface temperature may also influence photosynthesis and microbial activities. Relatively high temperature can stimulate primary productivity and microbial activities, thus affecting the production and consumption of OC. Primary productivity indexes such as phytoplankton concentration and net primary productivity were also involved in model building. Sedimentary properties, such as sediment mixed-layer thickness, sediment thickness and TOC content, were the main controlling factors of OC degradation and burial in sediments[26,60].

Detailed information for the data of all of these environment variables, including spatial resolution, time period and sources, is provided in Supplementary Table 1. To ensure the spatial correspondence of each environmental variable dataset, we resampled all of the datasets to match a 10′ × 10′ resolution.

Feature selection and machine learning models

A global dataset of 2,559 river particle observations and 1,325 coastal surface sediment observations (Extended Data Fig. 1a,b) was compiled to train and test machine learning models. A total of 28 and 15 environmental variables were selected, respectively, to build up reliable predictive machine learning models for river particles and coastal sediments. For machine learning models, more variables may not necessarily improve model performance, but may sometimes lead to poor model performance, unnecessary (and undesirable) high model complexity and uncertainty propagating therein[61]. To obtain the optimal assembly of explanatory variables, we used Pearson correlation analysis and the random forest method to filter out the most important variables. Specifically, correlations were examined between $\Delta^{14}C$ values and each environmental variable using SPSS, and important variables were selected based on correlation coefficients and significances. The randomForest R package was used to train the models with different predictor combinations by examining the tenfold cross-validation and electing the optimal combination of independent variables with the best agreement between predicted and observed $\Delta^{14}C$ values. The optimal assembly of explanatory variables was determined by combining these two feature selection methods. Eleven variables are included in the model for predicting $\Delta^{14}C$ values of riverine particles, including mean annual temperature, mean annual precipitation, elevation, slope, modelled sediment yield, *R* factor of rainfall erosivity, contents of clay and silt, soil OC content, cation exchange capacity, net primary productivity and gross domestic product. To predict $\Delta^{14}C$ values in coastal sediments, different sets of environmental variables were used for Arctic regions and non-Arctic regions. The optimal variables in the Arctic region include dissolved inorganic C, salinity, p_{CO_2}, flow velocity, tidal range, net primary productivity, phytoplankton concentration, sea subsurface temperature, seawater transparency, mixed-layer thickness and water depth. The optimal variables in the non-Arctic region include TOC content, flow velocity, phytoplankton concentration, salinity, tidal range, sea subsurface temperature, sea surface temperature, water depth, dissolved inorganic C, p_{CO_2} and sediment thickness.

To build up reliable models for predicting the $\Delta^{14}C$ values of riverine particles and coastal sediments, we compared the different approaches, including multivariable linear regression, *k*-nearest neighbour, decision tree, neural network, boosting, random forest and support vector machine (SVM). The multivariable linear regression method is mainly used to describe the linear relationship between explanatory variables with dependent variables. However, when the number of dependent variables is too large, the model can overfit or underfit. The *k*-nearest neighbour method is a nonparametric approach that assigns weights to distances based on sample proximity, to reduce the impact of outliers and to improve model robustness[25]. The decision tree is a hierarchical classifier that recursively partitions a dataset into increasingly homogenous subsets (referred to as nodes) to predict class membership[62]. However, deep decision trees with sparse leaf nodes may lead to overfitting, thus reducing the model's generalization

ability. The neural network is usually a set of neurons that connects the input layer, hidden layers and output layer, and the key parameters are the number of hidden layers and neurons[26].

Compared with the single model described above, ensemble models (including bagging, boosting and stacking algorithms) improve the model's generalization ability by amalgamating multiple models, thereby mitigating the risk of overfitting associated with a single model. The boosting method, such as Gradient Boosting and eXtreme Gradient Boosting, is currently the dominant tree-based ensemble learning algorithm due to its powerful and robust predictability[63,64]. The boosting algorithm enhances weak learners through iterative training on a dataset, adjusting their sample weights based on error rates until a set number of weak learners is attained. These are then amalgamated into a robust learner. Models such as Gradient Boosting and eXtreme Gradient Boosting, renowned for their efficacy in ensemble settings, mitigate overfitting by fine-tuning tree parameters, step size and penalty coefficients. In contrast, bagging diverges by randomly sampling the dataset for each weak learner's creation, fostering multiple independent learners via random sampling. Random forest advances bagging by employing decision trees as weak learners and incorporating random feature selection, thereby optimizing generalization and curbing overfitting through precise tree parameter adjustments. The SVM method operates under the assumption that the joint distribution of the input and output variables is unknown, yet a correlation between them exists[65]. The SVM method projects the input features into a higher-dimensional feature space using kernel functions, thereby transforming linearly inseparable features to separable ones and iteratively adjusting the hyperplane to find the optimum solution[66,67]. SVM can enhance prediction accuracy by utilizing the outputs of multiple models as inputs to train a new model through a stacking algorithm.

All compiled datasets were randomly divided into two parts, with 70% of the full dataset used for training and the remaining 30% used for testing. Traditional evaluation methods may lack accuracy when assessing a model's performance with limited test data, often due to the potential lack of representativeness. Cross-validation mitigates this by partitioning the data into multiple folds and iteratively training and validating the model across different folds. This approach yields a more reliable and consistent evaluation of performance, which further reduces the risk of overfitting or underfitting. A tenfold cross-validation method was used to evaluate the model performance. During this tenfold cross-validation process, data for model training were divided into ten subsets of the same size; every nine subsets of data were used for model training, with the one left used for model validation. To reduce the uncertainty associated with stochastic sampling and identify the most predictive models, we trained the machine learning models 100 times and optimized for the highest R^2. In each run, we employed three iterations of tenfold cross-validation for model training. In addition to ensuring an adequate sample size in the training dataset, employing ensemble models and performing cross-validation methods, suitable feature selection can also help to mitigate overfitting because it can reduce the model complexity and reduce the impact of irrelevant or redundant features. For $\Delta^{14}C$ in riverine particles and coastal sediments, we respectively constructed several models based on the compiled datasets, using the randomForest R package with feature selection and the MATLAB toolbox (the Statistics and Machine Learning Toolbox and Deep Learning Toolbox) without feature selection. We evaluated modelling performance by comparing the coefficient of determination (R^2) and mean absolute error value; higher R^2 and lower mean absolute error values represent better model performance. Among the seven models used, the SVM performed the best. Optimizing hyperparameters in the SVM, such as the regularization parameter C and the type of kernel function, can improve model effectiveness. Regularization reduces model complexity, prevents overfitting and improves generalization to new data. The kernel function in SVM addresses nonlinearity in data and the radial basis kernel function is often preferred because of its better simplicity and performance compared with linear and polynomial kernels in most cases. To identify the optimal combination of hyperparameter settings for the SVM, we conducted hyperparameter tuning using the tenfold cross-validation method for the environmental variables based on grid search and optimization algorithms. Ultimately, the particle swarm optimization SVM (PSO-SVM), based on biological optimization algorithms, showed the best performance in predicting $\Delta^{14}C$ values of riverine particles among a series of SVMs; the grid search method-based SVM (GSM-SVM) showed the best performance in predicting $\Delta^{14}C$ values, $\delta^{13}C$ values and TOC contents in coastal sediments. The GSM-SVM was therefore used to predict the $\Delta^{14}C$ values in global river particles, and GSM-SVM was used for $\Delta^{14}C$ values, $\delta^{13}C$ values and TOC contents in global coastal sediments.

Conversion from $\Delta^{14}C$ values to ^{14}C ages

Radiocarbon data are variably reported as $\Delta^{14}C$ values, fraction modern (F_m) values and/or radiocarbon ages (^{14}C ages). We used the following formulas[68] to convert $\Delta^{14}C$ values to ^{14}C ages:

$$\Delta^{14}C = 1,000 \times \left(F_m \times \exp^{-\lambda(y-1,950)} - 1\right) \quad (1)$$

$$F_m = \left(\left(\Delta^{14}C/1,000\right) + 1\right) \times \exp^{\lambda(y-1,950)} \quad (2)$$

$$^{14}C_{age} = -8,033 \times \ln\left[F_m\right] \quad (3)$$

whereby $\lambda = 1/8,267\ yr^{-1}$ and y is the year of sample collection and measurement. Notably, we assume identical years of collection and measurement because such information is typically not reported and because their minor difference (less than 5 years) does not introduce a significant error in the context of this study[17].

Global mapping of predictors

For riverine particles, the geographic location of prediction points is defined according to the classical ordering system in the attribute table of the HydroBASINS database[69]. In this classical ordering system, order 1 represents the main stem river from sink to source, order 2 represents all tributaries that flow into an order 1 river, order 3 represents all tributaries that flow into an order 2 river and so on.

First, we calculated the coordinate of the geometric centre of each sub-basin (corresponding to river order 7) in each river basin using the Calculate Geometry function in the ArcGIS attribute table. Then, we allocated the predicted riverine $\Delta^{14}C$ values at the sub-basin centre to the nearest river order 7 using the Near function in ArcToolbox. Finally, 22,442 data points were allocated across global river basins.

For coastal sediments, global maps with a consistent spatial resolution of 10′ × 10′ were generated using ArcGIS Pro, representing 99,807 sites of the coastal margins with water depths of no more than 200 m. First, the attributes of selected explanatory variables at each site were extracted in ArcGIS Pro and exported to the predictor database. Then, we fed the predictor database to the trained GSM-SVM to predict the gridded $\Delta^{14}C$ values, TOC contents and $\delta^{13}C$ values in global coastal sediments.

Data availability

Source data are provided with this paper in the Supplementary Information and archived at https://doi.org/10.6084/m9.figshare.24268657 (ref. 70).

Code availability

The code generated and/or used for the analyses in this study is archived at https://doi.org/10.6084/m9.figshare.24268657 (ref. 70).

References

54. Hilton, R. G. Climate regulates the erosional carbon export from the terrestrial biosphere. *Geomorphology* **277**, 118–132 (2017).
55. Syvitski, J. et al. Earth's sediment cycle during the Anthropocene. *Nat. Rev. Earth Env.* **3**, 179–196 (2022).
56. Bao, R. et al. Influence of hydrodynamic processes on the fate of sedimentary organic matter on continental margins. *Glob. Biogeochem. Cycles* **32**, 1420–1432 (2018).
57. Ji, R. B. et al. Influence of ocean freshening on shelf phytoplankton dynamics. *Geophys. Res. Lett.* **34**, L24607 (2007).
58. Gouveia, N. A. et al. The salinity structure of the Amazon River plume drives spatiotemporal variation of oceanic primary productivity. *J. Geophys. Res. Biogeosci.* **124**, 147–165 (2019).
59. Laruelle, G. G. et al. Continental shelves as a variable but increasing global sink for atmospheric carbon dioxide. *Nat. Commun.* **9**, 454 (2018).
60. Magill, C. et al. Transient hydrodynamic effects influence organic carbon signatures in marine sediments. *Nat. Commun.* **9**, 4690 (2018).
61. Xiang, D. F., Wang, G. S., Tian, J. & Li, W. Y. Global patterns and edaphic-climatic controls of soil carbon decomposition kinetics predicted from incubation experiments. *Nat. Commun.* **14**, 2171 (2023).
62. Hansen, M. C. et al. Monitoring conterminous United States (CONUS) land cover change with Web-Enabled Landsat Data (WELD). *Remote Sens. Environ.* **140**, 466–484 (2014).
63. Friedman, J. H. Greedy function approximation: a gradient boosting machine. *Ann. Stat.* **29**, 1189–1232 (2001).
64. Chen, S. C. et al. A high-resolution map of soil pH in China made by hybrid modelling of sparse soil data and environmental covariates and its implications for pollution. *Sci. Total Environ.* **655**, 273–283 (2019).
65. Tahmasebi, P., Kamrava, S., Bai, T. & Sahimi, M. Machine learning in geo- and environmental sciences: from small to large scale. *Adv. Water Resour.* **142**, 103619 (2020).
66. Noble, W. What is a support vector machine? *Nat. Biotechnol.* **24**, 1565–1567 (2006).
67. Chapelle, O. Training a support vector machine in the primal. *Neural Comput.* **19**, 1155–1178 (2007).
68. McNichol, A. P. & Aluwihare, L. I. The power of radiocarbon in biogeochemical studies of the marine carbon cycle: insights from studies of dissolved and particulate organic carbon (DOC and POC). *Chem. Rev.* **107**, 443–466 (2007).
69. Lehner, B. & Grill, G. Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. *Hydrol. Process.* **27**, 2171–2186 (2013).
70. Wang, C. et al. Global patterns of organic carbon transfer and accumulation across the land–ocean continuum constrained by radiocarbon data. *Figshare* https://doi.org/10.6084/m9.figshare.24268657 (2024).

Acknowledgements

This work was supported by the National Natural Science Foundation of China (grant numbers 42106056 and 42293261 to C.W. and Y.W.), China Postdoctoral Science Foundation (2021M691504 to C.W.), Fundamental Research Funds for the Central Universities (grant number 0209-14370407 to X.Z.) and Dutch Ministry of Education, Culture and Science through the Netherlands Earth System Science Centre (to J.W. and J.J.M.).

Author contributions

C.W. and X.Z. conceptualized the study. Y.Q. and C.W. performed the machine learning and other modelling approaches. Y.Q., C.W. and Z.H. performed the data analysis. C.W., Z.H. and C.Z. collected the data. Y.Q. performed the visualization. J.W., X.Z., J.J.M. and Y.W. contributed to the methodology of modelling and analyses. C.W., Z.H. and J.W. prepared the manuscript. All co-authors reviewed and commented on the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41561-024-01476-4.

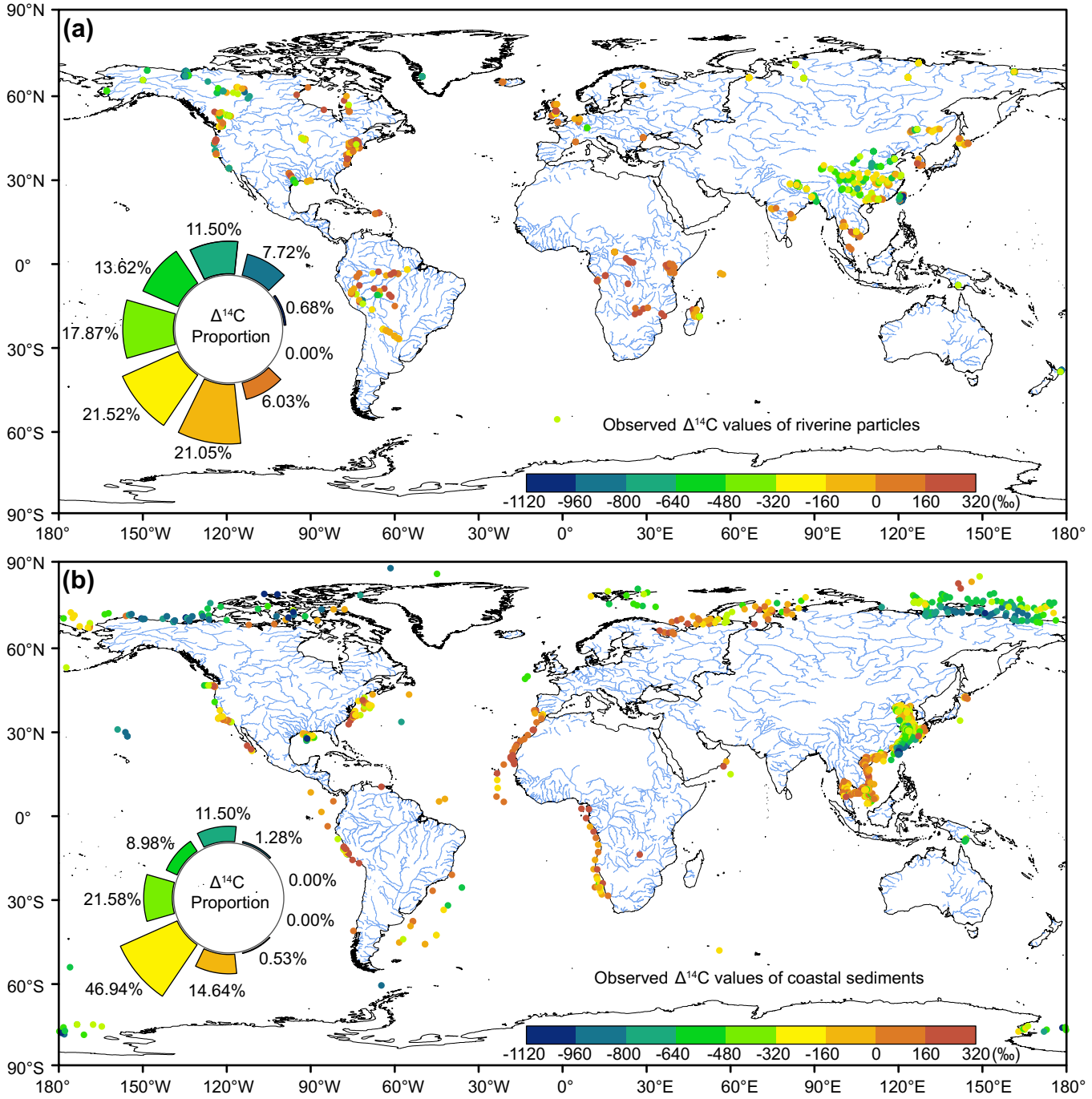
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41561-024-01476-4.

Correspondence and requests for materials should be addressed to Chenglong Wang or Xinqing Zou.

Peer review information *Nature Geoscience* thanks Elizabeth Canuel, Neal Blair and Ewa Burwicz-Galerne for their contribution to the peer review of this work. Primary Handling Editor: Xujia Jiang, in collaboration with the *Nature Geoscience* team.

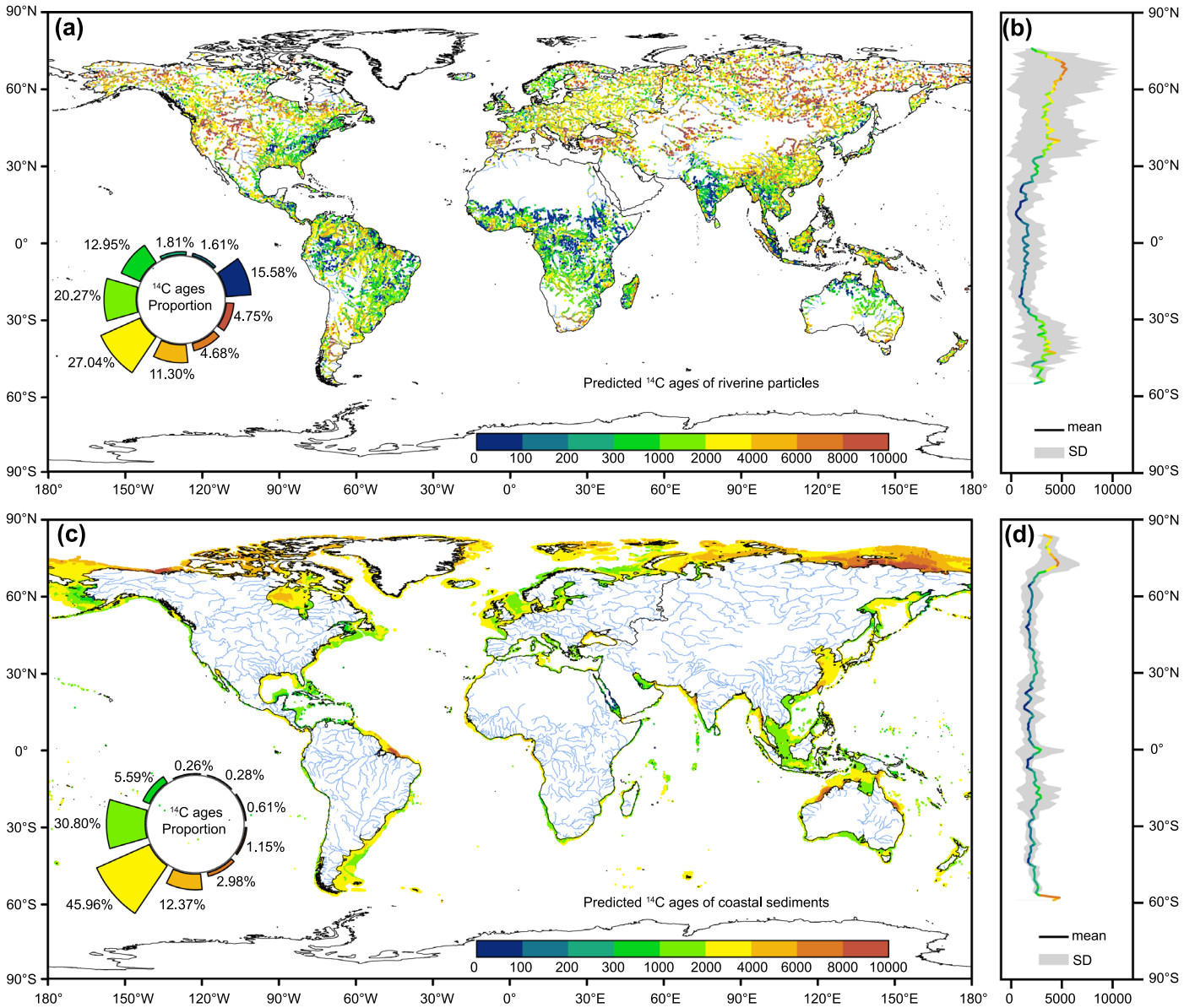
Reprints and permissions information is available at www.nature.com/reprints.

https://doi.org/10.1038/s41561-024-01476-4



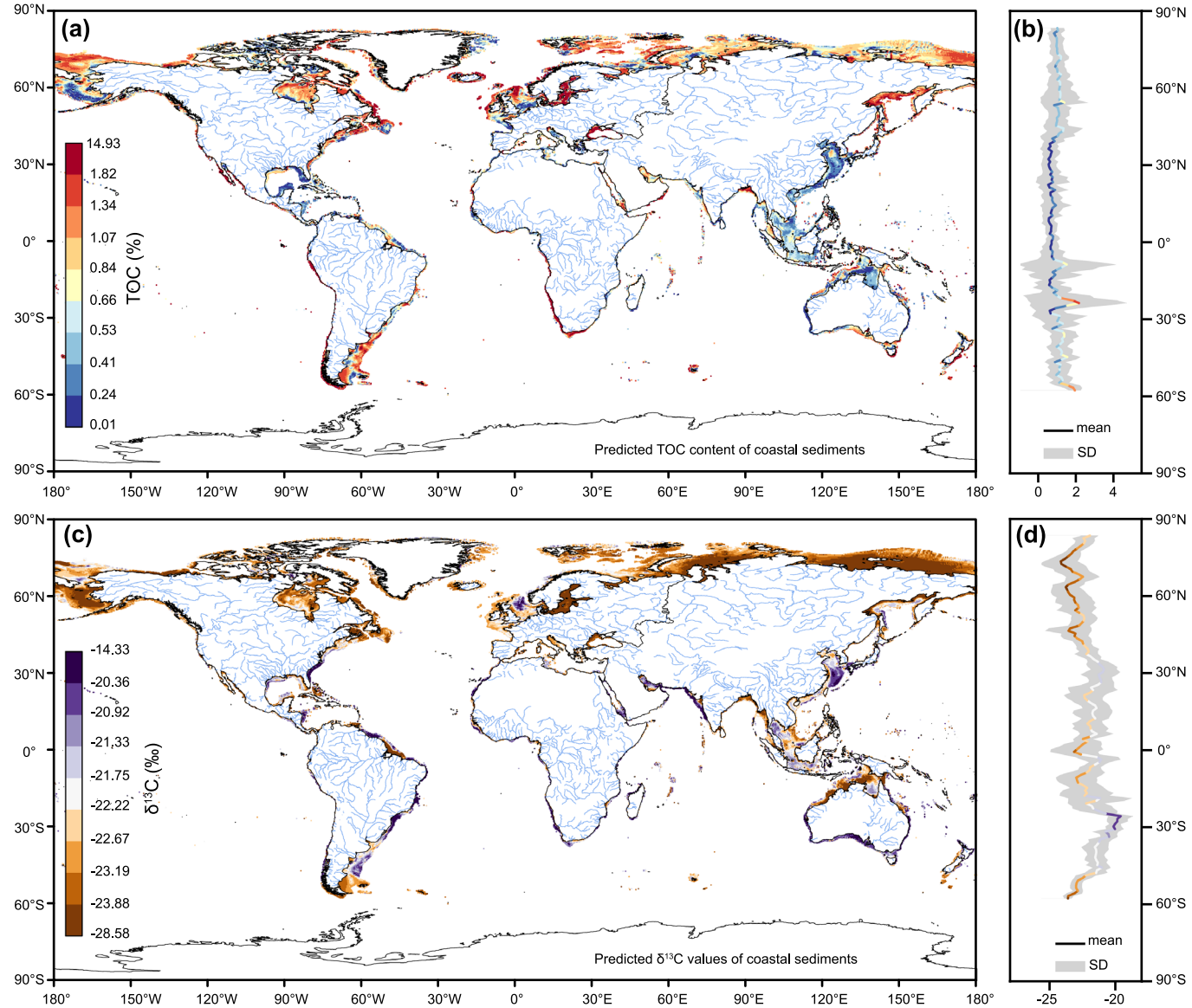
Extended Data Fig. 1 | Global distributions of $\Delta^{14}C$ observations in riverine particles and coastal sediments. (**a**) riverine particles (n = 2,559); (**b**) coastal sediments (n = 1,325).

https://doi.org/10.1038/s41561-024-01476-4



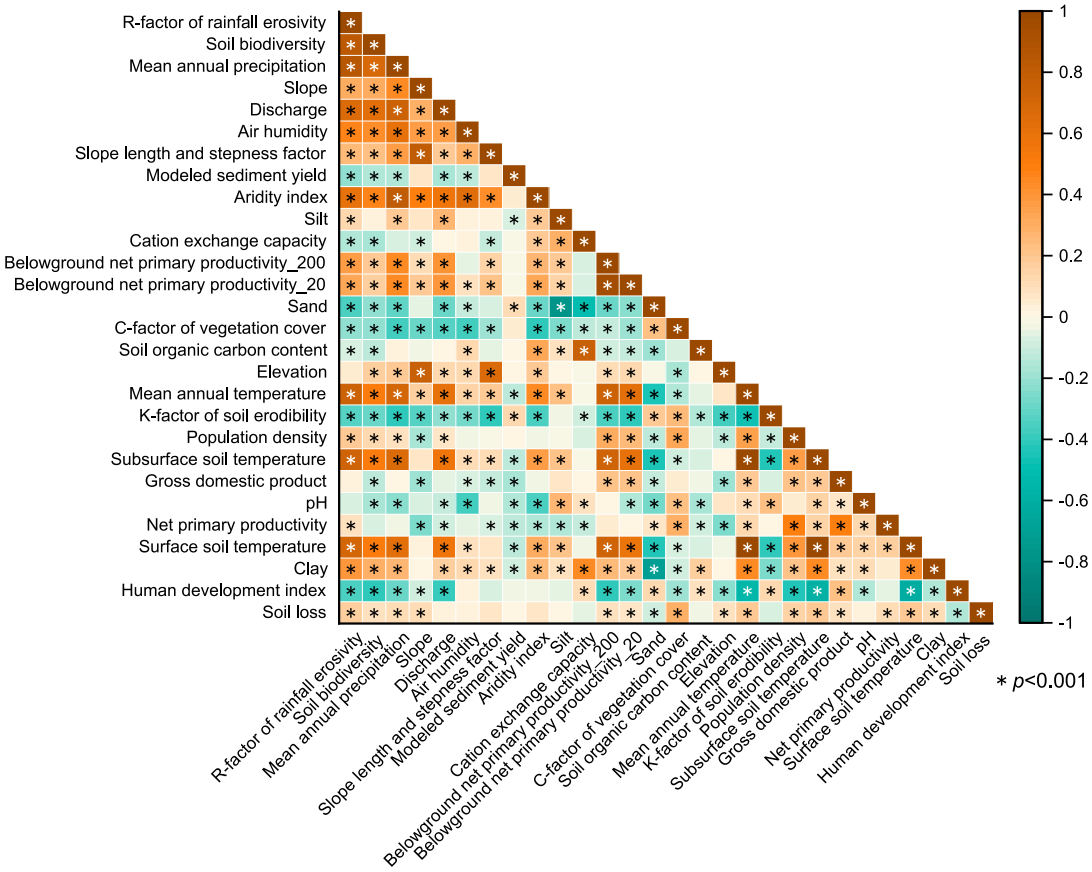
Extended Data Fig. 2 | Global distributions and latitudinal patterns of predicted ^{14}C ages in riverine particles and coastal sediments. a-d: (**a, c**) Global distributions and (**b, d**) latitudinal patterns of predicted ^{14}C ages in (**a, b**) riverine particles and (**c, d**) coastal sediments. The high-resolution global map of the predicted ^{14}C ages of riverine particles (**a**) is shown at river orders 1-7 defined in the 'Classical ordering system' of HydroBASINS (https://www.hydrosheds.org/products/hydrobasins) and includes 22,442 predicted values (Methods). The global map of the predicted ^{14}C ages of coastal sediments (**c**) is shown at a 10′ × 10′ spatial resolution and includes 99,807 predicted values (Methods). In panels (**b, d**), 'SD' means standard deviation. The predicted ^{14}C ages are converted from the predicted Δ^{14}C values in Fig. 2 following the approach in Methods.

https://doi.org/10.1038/s41561-024-01476-4



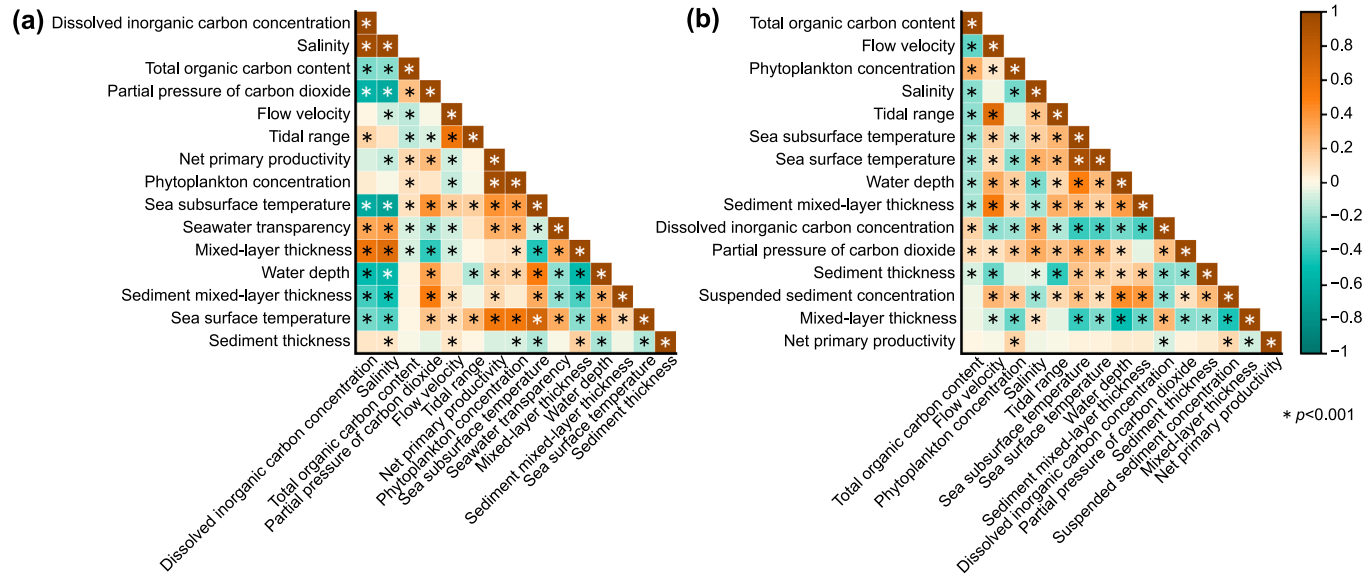
Extended Data Fig. 3 | Global distributions and latitudinal patterns of predicted TOC contents and $\delta^{13}C$ values in coastal sediments. a-d: (**a**, **c**) Global distributions and (**b**, **d**) latitudinal patterns of (**a**, **b**) predicted TOC contents and (**c**, **d**) $\delta^{13}C$ values in coastal sediments. The global maps of predicted TOC contents and $\delta^{13}C$ are shown at a 10′ × 10′ spatial resolution and include 99,807 predicted values. In panels (**b**, **d**), 'SD' means standard deviation.

https://doi.org/10.1038/s41561-024-01476-4



Extended Data Fig. 4 | The Pearson correlation coefficients between explanatory variables for riverine particles. '*' represents $p < 0.001$. 'Belowground NPP_200' represents belowground net primary productivity at 200-cm depth, and 'Belowground NPP_20' represents belowground net primary productivity at 20-cm depth.

https://doi.org/10.1038/s41561-024-01476-4



Extended Data Fig. 5 | The Pearson correlation coefficients between explanatory variables for coastal sediments. (**a**) Arctic regions and (**b**) non-Arctic regions. In the figures, '*' represents $p < 0.001$.