

Supporting Information

Follett et al. 10.1073/pnas.1407445111

SI Materials and Methods

Seawater was collected using wire-mounted Niskin bottles that had been cleaned and tested for radiocarbon contamination before use. Water was drained directly (without filtration) into clean polycarbonate bottles and frozen. Particulate organic carbon measured at Station ALOHA by the Hawaii Ocean Time-series program immediately before and after our sample collection shows 2.2–2.3 $\mu\text{M C}$ or <3% of TOC at 50 m, and 0.4 $\mu\text{M C}$ or <1% of TOC at 350 m, and we expect POC to represent <1% of TOC in our 500- and 2,000-m samples. Photo-oxidation of these small amounts of POC does not significantly affect our results. Samples were stored at -20°C until analysis.

In the laboratory, the concentration, $\delta^{13}\text{C}$, $\delta^{14}\text{C}$, and $\Delta^{14}\text{C}$ values were determined at each depth for a given sample. This measurement was done in conjunction with the National Ocean Sciences Accelerator Mass Spectrometry (NOSAMS) facility using a slightly modified version of their standard DOC oxidation protocol. Each bottle was acidified ($\text{pH} \leq 2.5$) using phosphoric acid and then split for use in a bulk analysis and in our time series analysis. The water for the bulk analysis was sparged with ultra-high purity helium gas for 1 h. At this point, it was irradiated with a 1,200-W, medium pressure mercury arc UV lamp (UV Doctor) for 4 h. A water cooling system is used to keep the pressure of the system below 800 torr. A shutter system was used to allow the lamp intensity to stabilize for 2.5 min before starting the irradiation. After this time the carbon dioxide generated from the oxidation was collected on a vacuum extraction line and quantified with a calibrated pressure gauge. We used the maximum sample volume possible (~ 1 L) for our measurements. The collected carbon dioxide was then cleaned and analyzed for $\delta^{13}\text{C}$ and $\Delta^{14}\text{C}$ by the NOSAMS staff. The $\delta^{13}\text{C}$ value was obtained on a separate gas split using isotope ratio mass spectrometry. Time series measurements were done in a similar fashion. For the first time point, the system was sealed, and the acidified sample was sparged with ultra-high purity helium gas for 1 h. After the 2.5-min warm-up time, the sample was irradiated for the proper time at which point the lamp was turned off and the carbon dioxide was collected and quantified. If the internal pressure neared 800 torr, then the cooling system was engaged. This threshold was reached after 8–12 min of oxidation. Each time step proceeded in this manner without additional 1-h sparges. An entire time series took 3 d to collect, and therefore the system was kept closed to the atmosphere with an overpressure of around 40 torr during the night hours. Data are shown in Table S1. The concentration data are corrected for blank mass by completing an entire serial oxidation experiment on seawater, which had been previously oxidized under UV light for 4 h to remove all DOC. The average blank was 0.35 μmol of carbon. However, when the splits were combined and recondensed, the total carbon recovered suggests an average blank between 0.1 and 0.2 μmol , consistent with the blanks obtained for standard bulk DOC radiocarbon measurements at NOSAMS. Individual time points contained too little carbon for reliable isotope measurements.

Serial UV-oxidation experiments on oceanic DOC have been reported by Beaupre and colleagues (1, 2) at Station M (Eastern Pacific). Unfortunately, $\delta^{13}\text{C}$ was not reported, making a direct comparison with our measurements impossible. One important difference between the two studies is the mass distribution across the time series. Beaupre et al. oxidize half of the material in the first time point, which leads to the apparent lack of fluctuations in the radiocarbon time series. Despite this, a simple two com-

ponent exponential model of their deep water data (1) fits quite well ($R^2 = 0.997$), suggesting that approximately 40% of the carbon in the deep ocean has a radiocarbon value below -720‰ and that two exponentials are sufficient to explain their mass time series. Using both $\delta^{14}\text{C}$ and $\delta^{13}\text{C}$ data provides information about additional DOC components. Mass only has logarithmic resolution in rate constant space. It is the fluctuations in $\delta^{14}\text{C}$ and $\delta^{13}\text{C}$ that allow us to access these intermediate rate constants. It should be noted here that, although $\Delta^{14}\text{C}$ is an appropriate correction for a single fractionation event, in general, isotopic fractionation is nonlinear, and the underlying distribution of potential rate constants and fractionation factors is best directly taken into account as done here.

It is important to our conclusions that the fluctuations we find in the isotopic time series are real and not artifacts of our analysis. Two types of errors could affect our time series: systematic (from blanks) and statistical errors. Large systematic errors appear insignificant to our $\delta^{14}\text{C}$ time series. The deviation between the bulk oxidation and mean value from the time series is less than 20‰ for all three depths which is consistent with the SD (23‰) for bulk samples taken from the same site and analyzed on the NOSAMS apparatus (3). Systematic errors due to a correlation between sample time and mass (1) were mitigated by choosing time points to collect approximately equal portions of the sample at each step. If the data trends seen here were due to systematic errors, we would expect to see the largest fluctuations in the deep water sample where concentrations are lower. The largest fluctuations are found in the surface water sample, contrary to this hypothesis. Correspondence in the structure of our time series between the surface, mid-depth, and deep waters suggests that statistical errors are small. To quantify this, we use the differences between ensuing data points in Table S1 (including the estimated values) to get an estimate of the local derivatives. We focus on the 500- and 2,000-m samples as their total concentration differs by less than 20% and they have a correlation coefficient of 0.82. A plot of both sets of derivatives corroborates that there are two inflection points and one local minima in both time series. We find that the derivatives differ between the two depths by an average value of 5‰ per time point and have an SD of 11‰ with a maximum deviation from the mean of 20‰. The perturbation errors of $\pm 10\text{‰}$ used in our Monte-Carlo analysis are consistent with the deviations found between these depths. Any larger statistical errors are inconsistent with the strong correspondence between these two time series.

The reproducibility of our time series was tested using two portions from the same 500-m station (Table S2). The samples were analyzed 2 y after the first suite of time series analysis, and the data may be influenced by changes in the UV lamp power and output. Systematic errors between the two time series were 0.5 μM , 0.6‰ $\delta^{13}\text{C}$, and 20‰ $\delta^{14}\text{C}$. These systematic shifts are small relative to the ranges in radiocarbon and stable carbon found in our analysis. The mean statistical errors were 0.3 μM , 1.6‰ $\delta^{13}\text{C}$, and 12.7‰ $\delta^{14}\text{C}$. These errors are consistent with the perturbation ranges used in our Monte-Carlo analysis of 0.4 μM , 1.2‰ $\delta^{13}\text{C}$, and 20‰ $\delta^{14}\text{C}$. Performing our data inversion on the first replicate and comparing it with the inversion from the 500-m sample used in the main text (Fig. S1) demonstrates the robustness of our distribution estimate. The second replicate was not used because of the missing stable isotope data. The larger errors in the distribution are most likely due to the missing time point 9 in the replicate time series and the poor resolution it created in the derivative there.

SI Text

DOC Photo-Oxidation Kinetics. The oxidation of DOC by UV radiation has long been considered a parallel first-order process (1, 4) such that

$$g(t) = \int_0^{\infty} g_0 p(k) e^{-kt} dk, \quad [\text{S1}]$$

where $g_0 = g(0)$ and the rate distribution $p(k) = \rho(k)/g(0)$. The time derivative

$$\dot{g} = -g_0 \int_0^{\infty} k p(k) e^{-kt} dk, \quad [\text{S2}]$$

is proportional to the initial concentration. For the parallel second-order case

$$g = \int_0^{\infty} \frac{g_0 p(k)}{1 + g_0 p(k) kt} dk, \quad [\text{S3}]$$

where

$$\dot{g} = \int_0^{\infty} \frac{-g_0^2 k p^2(k)}{[1 + g_0 p(k) kt]^2} dk. \quad [\text{S4}]$$

Initially, this derivative is proportional to the square of the concentration

$$\dot{g}|_{t=0} = -g_0^2 \int_0^{\infty} k p^2(k) dk. \quad [\text{S5}]$$

We can distinguish between the two cases by plotting the initial rate, $\dot{g}|_{t=0}$, vs. the scale factor, g_0 , or initial concentration. For a given sample of DOC at differing dilutions, varying g_0 , the initial rate should be well fit by a line through the origin

$$\dot{g}|_{t=0} = c_1 g_0, \quad [\text{S6}]$$

if the reaction is parallel first order (c_1 is a constant), and by a parabola with its vertex at the origin

$$\dot{g}|_{t=0} = c_2 g_0^2, \quad [\text{S7}]$$

if the reaction is parallel second order (c_2 is a constant).

We UV-oxidized seawater with a given DOC concentration and then repeated the experiment at different DOC dilutions. We plot the best estimate of the initial decay rate, based on the amount of DOC lost after 10 min of oxidation, vs. the initial concentration (Fig. S2). The data are much better fit by a line through the origin than by a parabola, suggesting that DOC degradation by UV light is a parallel first-order reaction rather than a second-order reaction as previously reported (2). Use of a finite time point to estimate the initial oxidation rate does not effect the linearity or the intercept for the first-order case and does not lead to linearity for the second-order case. Beaupré et al. (2) use the additional evidence that the decay of a pure organic compound, sucrose, is consistent with second-order kinetics while only consisting of a single molecule. The reaction conditions used in DOC oxidation experiments allow for a mechanistic explanation. Sucrose is unstable in water, and, under acidic conditions, hydrolyzes to fructose and glucose with a half-life between 1 and 10 h depending on temperature (5). It is

expected that UV oxidation of sucrose will follow parallel first-order kinetics as at least three compounds (sucrose, fructose, and glucose) would exist in solution.

Estimating the Isotopic Distribution. Estimating the isotopic distribution from our time series data requires a kinetic model (how compounds decay under UV light) and a method for assigning an isotope value to the mass at a given rate constant. Inverting for a continuous distribution of rates, a method that has been successfully applied to environmental degradation problems (6), is highly sensitive to noise in the time series data and is accomplished using methods of regularization. With the addition of isotopic data to the decay time series, these methods are not directly applicable, and thus the inherent sensitivity to noise was directly evaluated using a Monte-Carlo approach. We first simplify the continuous space, which has in principle an unlimited number of compound groups, into six discrete components. Six was chosen as the smallest number of components required to match the number of inflection points in our time series data, but not as a suggestion that only six compounds make up DOC. The isotope values, masses, rate constants, and system fractionation factors were then fit starting from a random initial condition. The best fit to the data was found using gradient search methods; multiple initial conditions were used to find the global best fit. This fitting procedure was performed for many realizations of the data, produced by random perturbations of the data within error bounds, and the probability distribution for mass as a function of isotope value was calculated from the many data fits. A mathematical presentation of this process follows.

As stated in the main text, the kinetic model for the ultraviolet oxidation experiment is

$$g_i(t) = \int_{-\infty}^{\infty} \rho_i(k) e^{-kt} dk. \quad [\text{S8}]$$

The function $g_i(t)$ is the amount of ^iC , where $i = 12, 13$, or 14 , remaining in the DOC sample after a time t ; $\rho_i(k)$ is the amount of ^iC associated with a first-order rate constant between k and $k + dk$. We seek $\rho_i(k)$ given $g_i(t)$. We use an n -component model to approximate $\rho_i(k)$

$$\rho_i(k) = \sum_{j=1}^n r_{ij} A_j D(k - \nu_{ij}), \quad [\text{S9}]$$

where A_j is the amount of material in the j th component, r_{ij} is its isotopic ratio, ν_{ij} is its decay constant, and D is the Dirac delta function. Combining Eqs. S8 and S9 yields

$$g_i(t) = \int_{-\infty}^{\infty} \left[\sum_{j=1}^n r_{ij} A_j D(k - \nu_{ij}) \right] e^{-kt} dk, \quad [\text{S10}]$$

which simplifies into the sum

$$g_i(t) = \sum_{j=1}^n r_{ij} A_j e^{-\nu_{ij} t}. \quad [\text{S11}]$$

The rate constants ν_{ij} can be further simplified by using an isotopic fractionation factor $\alpha_j \simeq 1$ where

$$\nu_{12,j} = \frac{\nu_{13,j}}{\alpha_j} = \frac{\nu_{14,j}}{\alpha_j^2}. \quad [\text{S12}]$$

Data from each experiment were collected at a set of m time points, t_l , where $l = 1, \dots, m$. We denote the mass of ^iC collected

at time t_i by $G_i(t_i)$. Isotopes other than ^{12}C make up less than 1% of carbon atoms and therefore $G_{12}(t_i)$ approximates the total mass to excellent approximation. Thus, in the experiment, we measure the mass collected, $G(t_i) = G_{12}(t_i)$, and the isotopic ratio of that material, $R_i(t_i)$, defined as

$$R_i(t_i) = \frac{G_i(t_i)}{G_{12}(t_i)}. \quad [\text{S13}]$$

We fit the data $G(t_i)$ and $R_i(t_i)$ with our model functions $\hat{R}_i(t_i, A_j, \nu_{ij}, r_{ij}, \alpha_j)$ and $\hat{G}(t_i, A_j, \nu_{ij})$. These functions are defined as

$$\hat{G}(t_i, A_j, \nu_{ij}) = g_{12}(t_{i-1}) - g_{12}(t_i), \quad [\text{S14}]$$

and

$$\hat{R}_i(t_i, A_j, \nu_{ij}, r_{ij}, \alpha_j) = \frac{g_i(t_{i-1}) - g_i(t_i)}{g_{12}(t_{i-1}) - g_{12}(t_i)}, \quad [\text{S15}]$$

to account for the discrete nature of the data sampling.

We seek the set of r_{ij} and A_j that, when used to compute $\hat{G}(t_i, A_j, \nu_{ij})$ and $\hat{R}_i(t_i, A_j, \nu_{ij}, r_{ij}, \alpha_j)$, best fits the data $G(t_i)$ and $R_i(t_i)$. This fit minimizes the function

$$F = \sum_{i=1}^n \frac{1}{N_{12}} \left[\frac{\hat{G}(t_i) - G(t_i)}{e_{12}(t_i)} \right]^2 + \frac{1}{N_{13}} \left[\frac{\hat{R}_{13}(t_i) - R_{13}(t_i)}{e_{13}(t_i)} \right]^2 + \frac{1}{N_{14}} \left[\frac{\hat{R}_{14}(t_i) - R_{14}(t_i)}{e_{14}(t_i)} \right]^2, \quad [\text{S16}]$$

where N_i is the number of data points for isotope i , and e_i is the error estimate for those measurements. For each exponential component, j , we have five parameters to fit: A_j , r_{2j} , r_{3j} , α_j , and ν_{12j} . We constrain $A_j \geq 0$, $0.96 \leq r_{2j} \leq 0.99$, $0 \leq r_{3j} \leq 1.15$, and $0.95 \leq \alpha_j \leq 1$. The fractionation factor was constrained based on prior laboratory experiments on organic compounds under UV light (4). We constrain the mean isotope values from the model to match the results from the bulk measurement. We have access to 31 data points in our experiment including the bulk measurements and require four to seven exponential components to match the time series, based on the number of turning points in $R_i(t_i)$. We used $n = 6$ exponential components. Because the solution is sensitive to any noise in the data (the problem is ill posed), we compile an ensemble of solutions and find their distribution in isotope space.

The computation of this distribution accounts for measurement errors as follows. We perform an analysis for 1,000 random perturbations of the data $G(t_i)$ and $R_i(t_i)$ within the bounds of $\pm e_i(t_i)$. For each dataset, we minimized F , starting from 10 different random initial guesses of the solution. From this set of 10, the best parameter set was kept. At the end of our calculations, we had a set of 1,000 solutions. The isotope distributions were estimated as histograms of these solutions.

Stable Isotope Distribution. Stable carbon $\delta^{13}\text{C}$ estimates spanned the entire allowable range with a tendency towards bulk average values with increasing depth (Fig. S3). This wide range supports the idea that external sources are quantitatively important to the DOC reservoir. The decreasing variability with depth is consistent with the decreasing range of isotope values measured in the time series. The range drops from approximately 30‰ in the surface to 25‰ at mid-depth and 20‰ in the deep ocean

sample. Without applying our model, either surface DOC is more isotopically diverse than deep water DOC or has a substantially higher average isotopic fractionation. As isotopic fractionation values are of similar size to the observed isotopic ranges in $\delta^{13}\text{C}$, fractionation is more relevant to this isotope, making the interpretation of our isotope distribution for $\delta^{13}\text{C}$ harder than for $\delta^{14}\text{C}$. However, both the time series data and the distribution of $\delta^{13}\text{C}$ within the DOC point to a DOC reservoir with diverse source material.

Radiocarbon Turnover Time. A population of DOC molecules has an age structure with both a set of discrete ages, a , and a mean age \bar{a} . The ages of the different chemical groups are reflected in the radiocarbon age, a_r , which is calculated from the measured isotopic ratio for the sample, R . The radiocarbon ratio, $r(a)$, of a sample of carbon with no material exchange changes as a function of its age, a , through radioactive decay as

$$r(a) = e^{-\lambda a}, \quad [\text{S17}]$$

where λ is the decay constant for ^{14}C and $r(0) = 1$. The radiocarbon age, a_r , is defined in terms of the measured isotopic ratio R as

$$a_r \equiv \frac{-\ln R}{\lambda}. \quad [\text{S18}]$$

When a reservoir of carbon contains a single, noncycling component, the radiocarbon age and the mean age are equivalent: $\bar{a} = a_r$. Once the carbon reservoir is cycling, Eq. S17 does not apply because a distribution of ages exist inside any given sample. In this case, the measured isotopic ratio, R , is an average isotopic ratio wherein $r(a)$ is weighted by the probability $P(a)$ that any given molecule has a certain age. Therefore

$$R = \int_0^{\infty} P(a)r(a)da. \quad [\text{S19}]$$

If a reservoir contains a single component in first-order steady state, then the age distribution is (7)

$$P(a) = \frac{1}{\tau} e^{-a/\tau}, \quad [\text{S20}]$$

where τ , the turnover time, equals \bar{a} . Substituting Eqs. S17 and S20 into Eq. S19 then yields

$$R = \int_0^{\infty} \frac{1}{\tau} e^{-a/\tau} e^{-\lambda a} da. \quad [\text{S21}]$$

Solving for the turnover time τ , we obtain

$$\tau = \frac{1 - R}{R\lambda}. \quad [\text{S22}]$$

Using Eq. S18, we find τ in terms of the radiocarbon age

$$\tau = \frac{e^{\lambda a_r} - 1}{\lambda}. \quad [\text{S23}]$$

Substituting $a_r = 12,000$ y, and $\lambda = 1/8,267$ y $^{-1}$, we find $\tau = 30,000$ y.

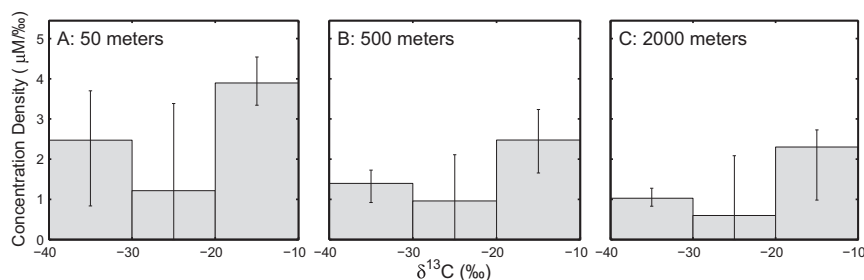


Fig. S3. Stable carbon distribution of DOC. The estimated isotopic distributions for 50- (A), 500- (B), and 2,000-m (C) depths. The area under each bar corresponds to the average concentration of the three evenly distributed fractions. Error bars are the 16% and 84% percentiles.

Table S1. Carbon mass and isotope data from the serial oxidation of DOC at Station ALOHA

Data source	UV time (min)	50 m				500 m				2,000 m			
		μM	$\delta^{13}\text{C}$	$\Delta^{14}\text{C}$	$\delta^{14}\text{C}$	μM	$\delta^{13}\text{C}$	$\Delta^{14}\text{C}$	$\delta^{14}\text{C}$	μM	$\delta^{13}\text{C}$	$\Delta^{14}\text{C}$	$\delta^{14}\text{C}$
Full oxidation	240.0	75.8	-21.3	-237.9	-232.2	45.2	-22.1	-395.7	-392.0	40.0	-22.7	-520.0	-517.7
Serial oxidation time series	1.5	5.1	-30.3	-224.5	-232.9	3.9	-27.5	-328.4	-331.8	2.9	-26.6	-471.4	-473.1
	3.6	6.5	-30.9	-189.9	-199.7	4.4	-29.5	-360.9	-366.8	2.8	-28.7	-517.9	-521.5
	6.3	7.1	-28.5	-175.4	-181.4	5.1	-29.1	-361.4	-366.7	3.3	-29.1	-528.6	-532.5
	9.9	7.9	-24.3	-179.3	-178.1	5.7	-27.0	-367.7	-370.4	4.0	-27.9	-527.1	-529.9
	15.4	8.2	-20.6*	-211.2*	-203.9*	6.5	-23.0	-378.4	-375.8	5.0	-20.8*	-534.7*	-530.7*
	24.1	8.5	-20.8	-252.8	-246.2	5.9	-19.7*	-405.0*	-398.5*	5.7	-20.8	-550.4	-546.5
	36.6	7.6	-22.4	-277.8	-273.9	4.5	-21.0	-433.7	-428.9	4.4	-19.8	-566.9	-562.2
	58.3	7.5	-24.3	-297.5	-296.5	3.8	-24.4	-443.5	-442.7	3.3	-24.2	-570.5	-569.7
91.9	6.1	-21.8	-319.8	-315.3	3.0	-25.6	-431.0	-431.7	2.4	-26.7	-549.6	-551.1	
240.0	8.0	2.1	-310.6	-271.8	3.8	-4.2	-448.0	-424.1	3.5	-10.4	-533.4	-519.2	
Bulk from time series	240.0	72.4	-21.6*	-244.2*	239.3*	46.5	-23.2*	-392.4*	-390.6*	37.4	-23.4*	-537.2*	-535.3*
North Pacific Subtropical Gyre (2, 3)			50 m				637 m				1,808 m		
		75	-20.9	-145	-138	42	-20.8*	-412	-407*	34	-20.7	-511	-507

Full oxidation data are from a duplicate sample that was completely oxidized by 4 h of UV irradiation before analysis. The concentration values were blank corrected. Analytic errors from the gas measurement and isotope analysis were less than $\pm 0.2 \mu\text{M}$, $\pm 0.1 \delta^{13}\text{C}$, and $\pm 11 \delta^{14}\text{C}$.

*Missing time series data (blanks in table) were interpolated to estimate the average isotope values. These estimates were not used in the time series analysis presented in the main text.

