

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15

# Probability distributions of radiocarbon in open compartmental systems

Ingrid Chanca<sup>1,2</sup>, Susan Trumbore<sup>1</sup>, Kita Macario<sup>2</sup>, Carlos A. Sierra<sup>1,3</sup>

<sup>1</sup>Max-Planck-Institut für Biogeochemie, Hans-Knöll-Str. 10, Jena, Germany.

<sup>2</sup>Laboratório de Radiocarbono, Instituto de Física, Universidade Federal Fluminense, Av. Litorânea s/n, Niterói, RJ, Brazil.

<sup>3</sup>Department of Ecology, Swedish University of Agricultural Sciences, Ulls väg 16, Uppsala, Sweden.

## Key Points:

- Predicted radiocarbon distributions in open compartments vary according to the year of observation and model structure
- Expected  $\Delta^{14}\text{C}$  values of ecosystem respiration are in accord with empirical  $\Delta^{14}\text{C}$  data
- Probability distributions of radiocarbon in open compartments provide insights into ecosystem dynamics

---

Corresponding author: Ingrid Chanca, [ichanca@bgc-jena.mpg.de](mailto:ichanca@bgc-jena.mpg.de)

16 **Abstract**

17 Radiocarbon ( $^{14}\text{C}$ ) is commonly used as a tracer of the carbon cycle to determine how fast  
 18 carbon moves between different reservoirs such as plants, soils, rivers or oceans. However  
 19 such studies mostly emphasize the mean value (as  $\Delta^{14}\text{C}$ ) of an unknown probability dis-  
 20 tribution. We introduce a novel algorithm to compute  $\Delta^{14}\text{C}$  distributions from knowledge  
 21 of the age distribution of carbon in compartmental systems at equilibrium. Our results  
 22 demonstrate that the shape of the distributions might differ according to the speed of cy-  
 23 cling of ecosystem compartments and their connectivity within the system, and are mostly  
 24 non-normal. The distributions are also sensitive to the variations of  $\Delta^{14}\text{C}$  in the atmosphere  
 25 over time, as influenced by the counteracting anthropogenic effects of fossil-fuel emissions  
 26 ( $^{14}\text{C}$ -free) and nuclear weapons testing (bomb  $^{14}\text{C}$ ). Lastly, we discuss insights that such  
 27 distributions can offer for sampling and design of experiments aiming to capture the precise  
 28 variability of  $\Delta^{14}\text{C}$  values in ecosystems.

29 **Plain Language Summary**

30 Radiocarbon is a radioactive isotope of carbon prominent in environmental sciences for  
 31 tracing the dynamics of ecosystems, especially as recent changes in atmospheric radiocar-  
 32 bon allow tracking excess  $^{14}\text{C}$  created by weapons testing in the atmosphere on timescales  
 33 shorter what than can be determined using radioactive decay. For climate changing mitiga-  
 34 tion, a crucial uncertainty is the time carbon captured through the photosynthesis spends  
 35 in ecosystems before being released. For this purpose, radiocarbon can be valuable as a  
 36 biological tracer; however, it is necessary to accurately link the real age of carbon and its  
 37 radiocarbon age, as they usually differ. Forests and soils systems are open systems, con-  
 38 necting components with intrinsically different cycling timescales, so that the mean age is  
 39 representing an age distribution that is not normally distributed. Here we developed an  
 40 algorithm to compute the  $^{14}\text{C}$  contents for models consisting of multiple interconnected car-  
 41 bon pools. Our approach, offers more accurate estimations of the mean  $^{14}\text{C}$  content of the  
 42 system and computations of the distribution of  $^{14}\text{C}$  within the system at different points in  
 43 time. From the results we can have more insights into the dynamics of the carbon cycle and  
 44 how to better design experiments to improve model-observations comparisons.

45 **1 Introduction**

46 Radiocarbon ( $^{14}\text{C}$ ) is a valuable tool for studying dynamical processes in living systems.  
 47 In particular, radiocarbon produced by nuclear bomb tests in the 1960s has been used in  
 48 many contexts as a tracer for the dynamics of carbon in different compartments of the  
 49 global carbon cycle, including the atmosphere, the terrestrial biosphere, and the oceans  
 50 (Goudriaan, 1992; Jain et al., 1997; Randerson et al., 2002; Naegler et al., 2006; Levin  
 51 et al., 2010). As a biological tracer, radiocarbon can be used to infer rates of carbon  
 52 cycling in specific compartments, and to infer transfers among interconnected compartments.  
 53 Therefore, radiocarbon is used as a diagnostic metric to assess the performance of models of  
 54 the carbon cycle (Graven et al., 2017), and new datasets are now emerging to incorporate  
 55 radiocarbon in model benchmarking (Lawrence et al., 2020).

56 Carbon cycling in biological systems can be represented using a particular class of  
 57 mathematical models called compartmental systems (Sierra et al., 2018). As carbon enters  
 58 a system such as the terrestrial biosphere, it is stored and transferred among a network  
 59 of interconnected compartments such as foliage, wood, roots, soils, and other organisms.  
 60 Compartmental systems represent the dynamics of carbon as it travels along the network of  
 61 compartments (Rasmussen et al., 2016; Sierra et al., 2018), and provides information about  
 62 the time carbon spends in particular compartments and the entire system (Rasmussen et  
 63 al., 2016; Sierra et al., 2017). Although there seems to be a direct relation between the time  
 64 carbon spends in a compartmental system and its radiocarbon dynamics, few studies relate  
 65 both concepts.

66 An open compartmental system contains inflows and outflows different from zero (Jacquez  
 67 & Simon, 1993). Timescales in open compartmental systems are usually characterized by  
 68 the concepts of *age* and *transit time* (Bolin & Rodhe, 1973; Rasmussen et al., 2016; Sierra  
 69 et al., 2017). In open systems such as the biosphere, the incorporation and release of car-  
 70 bon occurs continuously, but it is possible to define the concept of *age* as the time elapsed  
 71 since carbon enters the compartmental system until a generic time. The *transit time* can  
 72 be defined as the time the carbon needs to travel through the entire system, i.e., the time  
 73 elapsed between carbon entry until its exit.

74 In order to estimate these time metrics from  $^{14}\text{C}$  measurements, a model linking both  
 75 carbon and radiocarbon dynamics is required. Thompson and Randerson (1999) have used  
 76 impulse response functions from compartmental models to obtain ages, transit times, and  
 77 time-dependent radiocarbon dynamics. However, this approach is computationally expen-  
 78 sive and can introduce numerical errors if simulations are not long enough to cover the  
 79 dynamics of slow cycling pools.

80 Explicit formulas for age and transit time distributions in compartmental systems have  
 81 been recently developed (Metzler & Sierra, 2017). These formulas do not introduce nu-  
 82 merical errors and can describe entire age distributions of carbon for specific pools and  
 83 for the entire compartmental system. These age distributions suggest that radiocarbon in  
 84 compartmental systems may consist of a mix of different values, i.e., compartments could  
 85 be described in terms of radiocarbon distributions that relate the relative proportion of  
 86 carbon with a particular radiocarbon value. However, until now, radiocarbon is reported  
 87 and modeled as a single quantity, rather than the mean of an underlying distribution.

88 Knowledge of the distribution of  $^{14}\text{C}$  overlaid on the  $^{12}\text{C}$  distribution (C mass) in a  
 89 compartmental system might give important insights on the model structure that better fits  
 90 existing data. For example, by comparing the signature of radiocarbon in the pools and  
 91 their outfluxes, we get insights on the size of the pool model that describes the ecosystem.  
 92 Conversely, empirical knowledge of the radiocarbon distribution of a particular system, can  
 93 play a significant role in determining the most appropriate model to describe a system.

94 Model-data comparisons using radiocarbon are made more complex by the fact that  
 95 atmospheric  $^{14}\text{C}$  is continuously changing. This is particularly important after the 1960s  
 96 when the nuclear bomb tests liberated large amounts of thermal neutrons to the atmosphere,  
 97 contributing to the formation of radiocarbon (bomb  $^{14}\text{C}$ ). In addition, large quantities of  
 98 fossil-fuel derived carbon ( $^{14}\text{C}$ -free) have been emitted to the atmosphere, diluting the at-  
 99 mospheric radiocarbon signal and producing a fast decline of radiocarbon values in recent  
 100 years (Graven et al., 2017). Therefore, we would expect a different radiocarbon distribution  
 101 for every year in a compartmental system.

102 Obtaining a simple and accurate method to estimate radiocarbon distributions as a  
 103 function of the year of observation is, therefore, of great interest for experimental and  
 104 modeling studies.

105 The main objective of this manuscript is to introduce a method to obtain distributions  
 106 of radiocarbon in compartmental systems at steady-state. In particular, we ask the fol-  
 107 lowing research questions: (i) How do distributions of radiocarbon change over time as a  
 108 consequence of changes in atmospheric radiocarbon? (ii) How do empirical data compare  
 109 to these conceptual radiocarbon distributions? (iii) What insights can these distributions  
 110 provide for experimental and sampling design for improving model-data comparisons by  
 111 capturing the entire variability of  $\Delta^{14}\text{C}$  values?

112 The manuscript is organized as follows: First, we provide the necessary theoretical  
 113 background to obtain age and transit time distributions from compartmental systems. Sec-  
 114 ond, we describe an algorithm that computes radiocarbon distributions for particular years  
 115 using an age or a transit time distribution and an atmospheric radiocarbon curve. Third,  
 116 we present an application of our algorithm to a soil compartmental system addressing the

117 research questions above. Finally, we discuss our results in the context of other applications  
 118 and potential new insights from our approach.

## 119 2 Age and transit time distributions in compartmental systems

### 120 2.1 Compartmental systems

121 Compartmental systems describe the temporal dynamics of matter as it travels through  
 122 a network of compartments until its final release from the system. A set of compartments  
 123 is translated mathematically as a set of linear or non-linear ordinary differential equations  
 124 (ODE), whose solutions are the amount of matter in each compartment at a certain time.

125 We will consider here linear autonomous compartmental systems, characterized by the  
 126 mass of carbon at time  $t$  in  $m$  compartments as the vector  $\mathbf{x}(t)$ . The mass of carbon in the  
 127 compartments changes over time according to the following expression

$$128 \quad \frac{d\mathbf{x}(t)}{dt} = \dot{\mathbf{x}}(t) = \mathbf{u} + \mathbf{A} \mathbf{x}(t), \quad \mathbf{x}(t = 0) = \mathbf{x}_0, \quad (1)$$

129 where the vector  $\mathbf{u}$  represents the inputs of carbon into the system, and the  $m \times m$  com-  
 130 partmental matrix  $\mathbf{A}$  contains in its diagonal entries the cycling rates of the compartments,  
 131 while the off-diagonal entries consist of the transfer rates among them. In particular, the  
 132 compartmental matrix in most ecosystem carbon models has an internal structure reflecting  
 133 transfers between the components (coefficients  $\alpha_{i,j}$ , representing the proportion of C trans-  
 134 ferred from compartment  $j$  to compartment  $i$ ) and cycling rates  $k_i$  reflecting assumptions  
 135 of first-order kinetics of loss (at rate  $C_i k_i$ ) from any given compartment:

$$136 \quad \mathbf{A}_{m,m} = \begin{pmatrix} -k_1 & \alpha_{1,2}k_2 & \cdots & \alpha_{m,m}k_m \\ \alpha_{2,1}k_1 & -k_2 & \cdots & \alpha_{2,m}k_m \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m,1}k_1 & \alpha_{m,2}k_2 & \cdots & -k_m \end{pmatrix}, \quad (2)$$

137 This matrix contains information on the dynamics, structure, and size of a compartmen-  
 138 tal model. The rate of exit of carbon from the system can also be obtained from this matrix  
 139 by summing all column elements; i.e., the outputs from a pool that are not transferred to  
 140 other pools are assumed to leave the compartmental system.

141 The information of the amount of carbon entering the system to be partitioned among  
 142 the compartments is contained in the input vector

$$143 \quad \mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{pmatrix}. \quad (3)$$

144 Linear autonomous systems of the form of equation (1) have an equilibrium point or  
 145 steady-state solution  $\mathbf{x}^*$  given by

$$146 \quad \mathbf{x}^* = -\mathbf{A}^{-1} \mathbf{u}, \quad (4)$$

147 where the mass of the compartments do not change over time, and inputs are equal to  
 148 outputs for all compartments.

149 **2.2 Age distributions**

150 We define age  $\tau$  in a compartmental system as the time elapsed between the time of  
 151 carbon entry until some generic time (Sierra et al., 2017). For a time-independent system  
 152 in steady state, a probability distribution of ages of carbon in the compartments can be  
 153 obtained using stochastic methods. According to Metzler and Sierra (2017), the vector of  
 154 age densities for the compartments can be obtained as

155 
$$\mathbf{f}_a(\tau) = (\mathbf{X}^*)^{-1} \cdot e^{\tau \mathbf{A}} \cdot \mathbf{u} \tag{5}$$

156 where  $\mathbf{X}^* = \text{diag}(x_1^*, x_2^*, \dots, x_m^*)$  is the diagonal matrix with the steady-state vector of  
 157 carbon stocks as components, and  $e^{\tau \mathbf{A}}$  is the matrix exponential.

158 For the whole system, the age distribution is given by

159 
$$f_A(\tau) = -\mathbf{1}^\top \cdot \mathbf{A} \cdot e^{\tau \mathbf{A}} \cdot \frac{\mathbf{x}^*}{\|\mathbf{x}^*\|}, \tag{6}$$

160 where the symbol  $\|\cdot\|$  represents the sum of the masses in a vector.

161 **2.3 Transit Time distributions**

162 We define transit time as the time elapsed since carbon enters the compartmental  
 163 system until it leaves the boundaries of the system (Sierra et al., 2017). The transit time  
 164 is equivalent, therefore, to the age of the outflux. Metzler and Sierra (2017) also provide  
 165 an explicit formula to obtain the transit time density distribution for a time-independent  
 166 system at steady state as

167 
$$f_T(\tau) = -\mathbf{1}^\top \cdot \mathbf{A} \cdot e^{\tau \mathbf{A}} \cdot \frac{\mathbf{u}}{\|\mathbf{u}\|}. \tag{7}$$

168 These distributions are densities, so they integrate to 1

169 
$$\int_0^\infty f_A(\tau) d\tau = \int_0^\infty f_T(\tau) d\tau = 1. \tag{8}$$

170 **3 Methods**

171 **3.1 Radiocarbon distributions from age and transit time distributions**

172 We developed an algorithm to convert age and transit time distributions into  $\Delta^{14}\text{C}$   
 173 distributions for any given year of observation.

174 The algorithm works in three main steps, 1) homogenization, 2) discretization, and 3)  
 175 aggregation (Figure 1). We describe these three steps in detail in the sections below, using  
 176 mathematical notation for the system age distribution, but computations are similar for the  
 177 transit time distribution, and the age distribution of individual compartments.

178 **3.1.1 Homogenization of input data**

179 The main inputs for the algorithm are an age distribution  $f_A(\tau)$ , and an atmospheric  
 180 radiocarbon curve  $F_a(t)$  that provides the  $\Delta^{14}\text{C}$  value of atmospheric  $\text{CO}_2$  for a calendar  
 181 year  $t$ . To homogenize the time scales of both  $f_A(\tau)$  and  $F_a(t)$ , we define the year of  
 182 observation  $t_0$ , as the year of interest to produce the radiocarbon distribution.

183 Since we are interested in determining the radiocarbon values of material observed in  
 184 the system at time  $t_0$ , we will look in the radiocarbon curve  $-t$  years in the past to obtain

185 the radiocarbon values in the system with an age  $\tau$ . Therefore, atmospheric radiocarbon  
 186 can be expressed as a function of age, i.e.,  $F_a(t_0 - t) = F_a(\tau)$  (Figure 1). Now, both the  
 187 system age distribution  $f_A(\tau)$  and the atmospheric radiocarbon curve  $F_a(\tau)$  are functions  
 188 of the continuous variable  $\tau$  that represents age.

189 Several atmospheric radiocarbon datasets can be found in the literature (Reimer et al.,  
 190 2013, 2020; Hogg et al., 2013, 2020; Hua et al., 2013; Levin et al., 1980; Levin & Kromer,  
 191 1997; Levin et al., 2010; Graven et al., 2017). Also forecasts of radiocarbon content in the  
 192 atmosphere can be found in the recent literature (Graven, 2015; Sierra, 2018). However,  
 193 these atmospheric radiocarbon datasets do not necessarily have the same resolution in time.  
 194 Some of them provide predictions or data at an annual or four-monthly time step, while  
 195 in other datasets, some ranges are spaced by decades. To homogenize the resolution of the  
 196  $\Delta^{14}\text{C}$  and to transform these radiocarbon datasets into a continuous function of  $\tau$ , we use  
 197 a cubic spline interpolation to obtain  $\Delta^{14}\text{C}$  values for any value of  $\tau$ . After this step,  $f_A(\tau)$   
 198 can be computed for any value of  $\tau \in [0, \infty)$ , and  $F_a(\tau)$  until the last available date in the  
 199 chosen radiocarbon atmospheric dataset.

### 200 3.1.2 Discretization

201 Although we have now the age distribution and the radiocarbon data as continuous  
 202 functions of age, we need to discretize these functions in intervals of size  $h$ . The reason  
 203 for this discretization is that the probability density function of age  $f_A(\tau)$  is a measure of  
 204 the relative likelihood of an infinitesimal amount of mass having an age  $\tau$ . But ultimately,  
 205 we are interested in the probability that a small mass has certain radiocarbon distribution.  
 206 Therefore, we need to discretize the probability density function to a probability mass  
 207 function along a discrete variable  $T \in [0, T_{\max}]$ . The new discrete probability function of  
 208 ages can be defined as

$$209 P_A(\tau \leq T \leq \tau + h) = \int_{\tau}^{\tau+h} f_A(\tau) d\tau. \quad (9)$$

210 From this probability function, we can compute the proportion of total mass in the  
 211 system with an age  $T$  as

$$212 M(T) = \|\mathbf{x}^*\| \cdot P_A(T), \quad (10)$$

213 where

$$214 \sum_0^{T_{\max}} P_A(T) \approx 1, \quad (11)$$

$$\sum_0^{T_{\max}} M(T) \approx \|\mathbf{x}^*\|.$$

215 Equation (11) implies that there is an approximation error by discretizing the contin-  
 216 uous density function to a finite set of discrete intervals. This approximation error can be  
 217 minimized by decreasing the size of the intervals  $h$  and extending  $T_{\max}$  as far as possible.

218 Once we discretize  $f_A(\tau)$  to  $P_A(T)$  and obtain discrete proportions of mass with certain  
 219 age  $M(T)$ , we proceed to discretize the atmospheric radiocarbon curve with respect to the  
 220 same discrete interval of ages  $T \in [0, T_{\max}]$ . This is simply done by computing  $F_a(\tau = T)$ ,  
 221 which makes the assumption that within each interval  $[\tau, \tau+h]$ , the atmospheric radiocarbon  
 222 value is equal to  $F_a(\tau)$ .

### 223 **3.1.3 Aggregation**

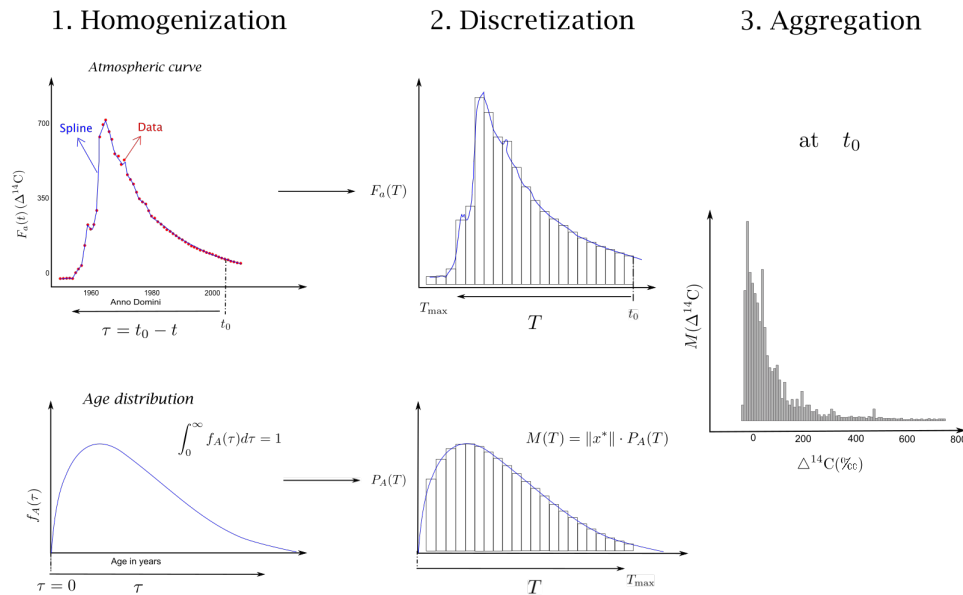
224 Now we are ready to combine the distribution of mass in the system at discrete age  
 225 intervals with the atmospheric radiocarbon curve. To do so, we first find for each value of  
 226  $T \in [0, T_{\max}]$  the corresponding values of mass  $M(T)$  and radiocarbon  $F_a(T)$ . Then, we  
 227 sum all the masses with similar  $\Delta^{14}\text{C}$  values. The result can be organized as the amount of  
 228 mass in discrete intervals of  $\Delta^{14}\text{C}$ ; i.e.,  $M(\Delta^{14}\text{C}) = M(F_a(T))$ .

229 These steps can also be visualized through the graphs in Figure 1.

230 We implemented these three steps in the R programming language, and use the package  
 231 SoilR (Sierra, Müller, et al., 2012) to obtain the age distribution of the pools, the whole  
 232 system, and the output flux (equivalent to the transit time) based on equations (5), (6), and  
 233 (7). The versions used here were R version 4.0.3 and SoilR version 1.1 (Sierra et al., 2014).

234 Since atmospheric  $^{14}\text{C}$  concentration for the past 55,000 years is principally known from  
 235 the radiocarbon curves, we could easily convert age into atmospheric  $\Delta^{14}\text{C}$ . By matching  
 236 the  $\Delta^{14}\text{C}$ -based-on-age values with the previously estimated densities, we built barplots,  
 237 gaining insight into the radiocarbon distributions for the model studied in this work. In the  
 238 algorithm we defined four functions: *PoolRDC*, *SystemRDC*, *TTRDC*, and *C14hist*. The  
 239 first three functions take the densities outputs, i.e., the carbon contents discretized by age,  
 240 from built-in SoilR functions, such as *transitTime* and *systemAge*. The densities are subset  
 241 to build bins through the *C14hist* function. The logical statements used to construct the  
 242 bins are based on the atmospheric  $\Delta^{14}\text{C}$  data and according to user-defined bin size  $b$ . This  
 243 structure allows one to plot histogram-like graphs, where the height of the bars represent  
 244 the amount of mass with corresponding  $\Delta^{14}\text{C}$  values. Thus, our algorithm initialize from a  
 245 compartmental matrix, an input vector and a radiocarbon calibration curve, and returns an  
 246 object containing masses of C and their matching decay-corrected  $\Delta^{14}\text{C}$  values, estimated  
 247 for any given observation year . The match is done by assuming the year of observation as  
 248 equivalent to the age of the pool or the system equals zero ( $t_0 - t = \tau = 0$ ). This means that  
 249 past years, or older pool or system ages, are equivalent to the  $\Delta^{14}\text{C}$  signal of the atmosphere  
 250 of those years corrected by the radioactive decay of  $^{14}\text{C}$  (average lifetime of 8,267 years, i.e.,  
 251 half-life of 5,730 years).

252 Besides the radiocarbon distributions for pools, whole system and output flux, one can  
 253 also compute the expected value of  $\Delta^{14}\text{C}$  from these distributions in any given observation  
 254 year. This is done by computing the mean of  $\Delta^{14}\text{C}$  weighted by the amount of carbon in  
 255  $\Delta^{14}\text{C}$  bins of size  $b$ . The standard deviation of the distribution is obtained as the square  
 256 root of the difference between the square of the expected value and the expected value of  
 257 the squares of  $\Delta^{14}\text{C}$  values.



**Figure 1.** Graphical visualization of the three main steps for the computation of radiocarbon distributions in a compartmental system using an atmospheric radiocarbon curve of the carbon inputs to the systems, and the age distribution of carbon in a compartmental system. Details about each step are provided in the main text.

258

### 3.2 Carbon Cycle models

259

Our approach can be used to obtain radiocarbon distributions for linear compartmental models of any size representing carbon cycling processes at different scales and for different biological systems.

260

261

262

We will focus here on a model that represents the dynamics of soil organic carbon at a temperate forest, which we call therein the Harvard Forest Soil (HFS) model. The model is based on measurements conducted at the Harvard Forest in Massachusetts, USA (Gaudinski et al., 2000; Sierra, Trumbore, et al., 2012). Soil samples collected in O-horizon, corresponding to 0 – 8 cm depth, and A-horizon (8 – 15 cm depth) were fractionated into seven soil fractions called: Dead Roots,  $O_i$ ,  $O_{e/a} L$ ,  $O_{e/a} H$ , A, LF ( $> 80 \mu\text{m}$ ), A, LF ( $< 80 \mu\text{m}$ ), and Mineral Associated. They were obtained as follows: The O-horizon was subdivided after hand-picking into leaf litter ( $O_i$  fraction), recognizable root litter ( $O_{e/a} L$  fraction) and humified, i.e., organic matter that has been transformed by microbial action, corresponding to the fraction  $O_{e/a} H$ . Samples from the A-horizon were fractionated by density into low-density and high-density portions. The high-density portion corresponds to the *Mineral Associated* fraction. The low-density portion is further subdivided by sieving into recognizable leaf larger than  $80 \mu\text{m}$  (A, LF ( $> 80 \mu\text{m}$ )) fraction) and smaller than  $80 \mu\text{m}$  (A, LF ( $< 80 \mu\text{m}$ )) fraction). Details about the methods employed to fractionate the samples can be found in Gaudinski et al. (2000).

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

The compartmental model consists of seven pools (Figure 2); one pool corresponds to dead roots  $x_1$ , and three pools correspond to three different types of organic matter in the surface layer (O) called  $O_i$ ,  $O_{e/a} L$ , and  $O_{e/a} H$ , which corresponds to pools  $x_2$ ,  $x_3$ , and  $x_4$  in the model. Two additional pools, called A, LF ( $> 80 \mu\text{m}$ ), representing material from the A horizon that floats in a dense ( $1 \text{ g cm}^{-3}$ ) liquid and does not pass through an  $80 \mu\text{m}$  sieve and A, LF ( $< 80 \mu\text{m}$ ) (low density fraction passing the sieve), represent the dynamics

278

279

280

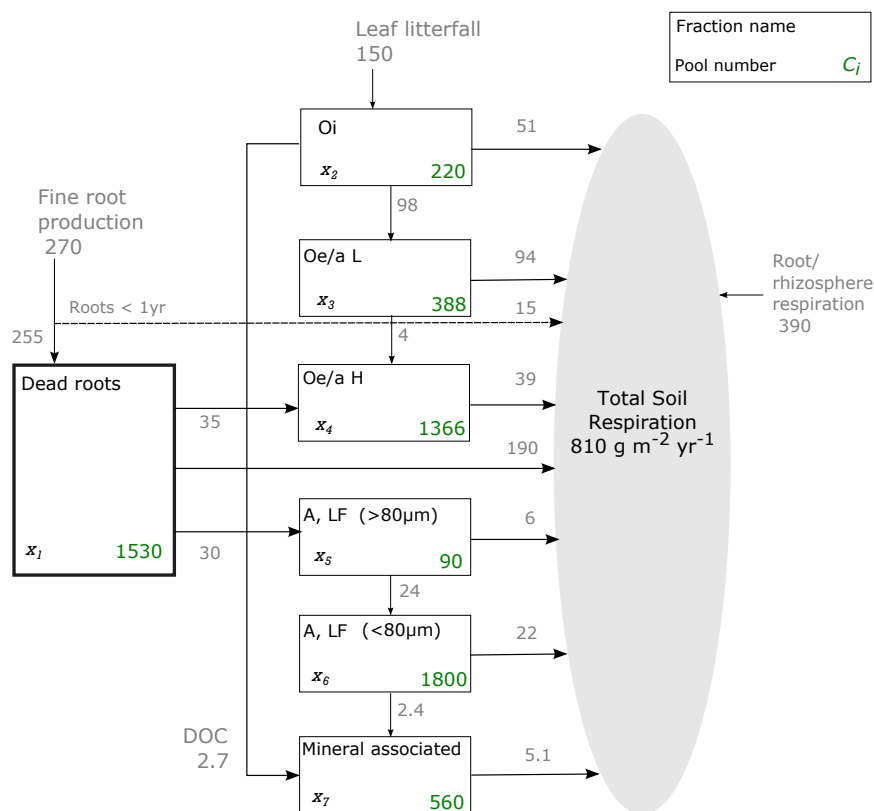
281

282



283 of two fractions in the soil A horizon with different granulometry,  $x_5$  and  $x_6$ , respectively.  
 284 The seventh pool  $x_7$  represents the dynamics of the mineral associated fraction (Sierra,  
 285 Trumbore, et al., 2012).

286 The HFS model was built by fitting of empirical radiocarbon data from the above  
 287 described samples. Details about the use of the data to build the compartmental model are  
 288 presented in Sierra, Trumbore, et al. (2012). For the same sites, there are independent data  
 289 (i.e., data not used for estimating the compartmental matrix) available. The independent  
 290 data used in this work consists of  $\Delta^{14}\text{C}$  measurements on total soil  $\text{CO}_2$  efflux collected  
 291 in the years 1996, 1998, 2002, and 2008. The number of samples measured corresponding  
 292 to the respective years was  $n = 12$ ,  $n = 28$ ,  $n = 23$ , and  $n = 10$ . We used these data to  
 293 compare the representativity of the mean  $\Delta^{14}\text{C}$  measurements to the expected  $\Delta^{14}\text{C}$  values  
 294 obtained through our algorithm.



**Figure 2.** Scheme of HFS model stocks ( $C_i$ ) and fluxes among compartments (adapted from Sierra, Trumbore, et al. (2012)).

295 The system of ODE for the HFS model can then be expressed in compartmental form  
 296 as

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \\ \dot{x}_6 \\ \dot{x}_7 \end{pmatrix} = \begin{pmatrix} 255 \\ 150 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -255/1530 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -150/220 & 0 & 0 & 0 & 0 & 0 \\ 0 & 98/152 & -98/388 & 0 & 0 & 0 & 0 \\ 35/255 & 0 & 4/98 & -39/1366 & 0 & 0 & 0 \\ 30/255 & 0 & 0 & 0 & -30/90 & 0 & 0 \\ 0 & 0 & 0 & 0 & 24/30 & -24/1800 & 0 \\ 0 & 3/152 & 0 & 0 & 0 & 3/25 & -5/560 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix}. \tag{12}$$

### 298 3.3 Set of parameters

299 As described before, in order to estimate the radiocarbon distributions and expected  
 300 values of  $\Delta^{14}\text{C}$ , the algorithm needs the following arguments: a compartmental matrix  $\mathbf{A}$ ,  
 301 containing the decomposition and transfer rates within the pools; an input vector  $\mathbf{u}$   
 302 containing the input mass to be partitioned among the compartments; the year of observation  
 303 (equivalent to year of *sampling* in an experimental framework); the number of years in the  
 304 past one aims to compute the distributions for; and a set of radiocarbon values in the atmo-  
 305 sphere, comprising the year of observation and the number of years chosen. An additional  
 306 argument is  $h$ , the discretization size described above, which has a default value of 0.1 years,  
 307 but could be modified according to user preferences.

308 For the HFS model,  $\mathbf{A}$  is the matrix in equation (12), with the form of equation (2),  
 309 and  $\mathbf{u}$  is the numeric vector in the same equation, with similar form as equation (3). We  
 310 estimated the radiocarbon distributions for different years of observation, in order to address  
 311 different research questions raised in this work. In the results we present the distributions  
 312 for the individual pools, total outflux and whole system, for the years: 1965, 2027 and  
 313 2100. Additionally, in the *Supplementary Material* we provide the non-stacked radiocarbon  
 314 distributions of individual pools, total outflux and whole system for the years 1950, 1965,  
 315 2027, and 2100. Radiocarbon distributions of the *outflux* are presented for the years: 1996,  
 316 1998, 2002, and 2008, as for those years we also have independent  $\Delta^{14}\text{C}$  data from soil  
 317  $\text{CO}_2$  efflux to compare to our estimations. For all those estimations, the number of years  
 318 of computation was 1,000 years. The bin size  $b$  ([‰]) for plotting the histograms was set as  
 319 10 for most of the radiocarbon distributions, except for the year 1965, where it was set up  
 320 to 40.

#### 321 3.3.1 Radiocarbon datasets

322 The radiocarbon values used for years in the past, e.g., AD 1965, were obtained by  
 323 merging the recently released IntCal20 calibration curve (Reimer et al., 2020), which com-  
 324 bines radiocarbon data and Bayesian statistical interpolation for the range 55,000 – 0 cal  
 325 BP (BP = *before present* = AD 1950), and the records of atmospheric radiocarbon data  
 326 compiled by Graven et al. (2017), from 1950 to 2015. Graven et al. (2017) also provides  
 327 radiocarbon data in one-year resolution on the range 1850 to 1949. However, since in this  
 328 range the estimations were partially based on the previous Northern Hemisphere calibration  
 329 curve (IntCal13, Reimer et al. (2013)), we decided to subset Graven et al. (2017)’s dataset  
 330 starting in AD 1950.

331 For the years in the future, such as AD 2027 and 2100, we made use of the forecast  
 332 simulations computed by Graven (2015), who simulated  $\Delta^{14}\text{C}$  values in the atmosphere for  
 333 four Represent Concentration Pathways of fossil fuel emissions: RCP2.6, RCP4.5, RCP6  
 334 and RCP8.5. In this work we use the predictions based on the high emissions scenario  
 335 (RCP8.5), starting in AD 2016.

336 The  $\Delta^{14}\text{C}$  values in all datasets used in this work are written as the deviation from the  
 337 standard representing the pre-industrial atmospheric  $^{14}\text{C}$  concentration. The raw published  
 338 values are already corrected for fractionation and decay with respect to the standard. It is  
 339 equivalent to  $\Delta$  in Stuiver and Polach (1977). Thus, the equation it follows is

$$340 \quad \Delta^{14}\text{C} = \left[ F^{14}\text{C} e^{\lambda_C(1950-y)} - 1 \right] \times 1000 \text{ [‰]} \quad (13)$$

341 where  $F^{14}\text{C}$  is the Fraction Modern ( $A_{SN}/A_{ON}$ ), i.e., the sample ratio normalised to  $\delta^{13}\text{C}$   
 342 by oxalic acid standard (OXII),  $\lambda_C$  is the updated  $^{14}\text{C}$  decay constant (equals  $1/8267 \text{ [y}^{-1}\text{)]}$ ,  
 343 and  $y$  is the year of measurement.

## 344 4 Results

### 345 4.1 Shape of the radiocarbon distributions and their change over time

346 Overall, our results show that even though the age and transit time distributions for  
 347 this compartmental system are static (Figure 3), the radiocarbon distributions are highly  
 348 dynamic. They change dramatically over time as the atmospheric  $\text{CO}_2$  source is affected  
 349 by the bomb spike and the Suess effect (Suess, 1955), i.e., the effect of the dilution of  
 350 radiocarbon in the atmosphere due to the emission of fossil fuels ( $^{14}\text{C}$ -free). Pools that cycle  
 351 fast, i.e., pools with sharp age distributions peaks, such as *Dead Roots* and *Oi*, followed most  
 352 closely the radiocarbon dynamics in the atmosphere, while pools that cycle slowly showed  
 353 a wide range of values. Consequently, the expected  $\Delta^{14}\text{C}$  values also vary largely.

354 The distributions we obtained for the compartments of the HFS model show very dif-  
 355 ferent shapes for the different compartments (Figures 4, 5, and 6, and Figures S2, S3, S4  
 356 and S5 in Supplementary Material). In 1965, just after the peak of bomb  $^{14}\text{C}$  in the atmo-  
 357 sphere due to nuclear weapons tests, pools that cycle fast had a wide  $\Delta^{14}\text{C}$  range with high  
 358 probability, due to the incorporation of radiocarbon values that changed rapidly over the  
 359 period AD 1950 – 1965. Compartments that cycle slowly have a narrower distribution with  
 360 their modes corresponding to negative  $\Delta^{14}\text{C}$  values, as they represent pre-bomb atmospheric  
 361 signals that varied less.

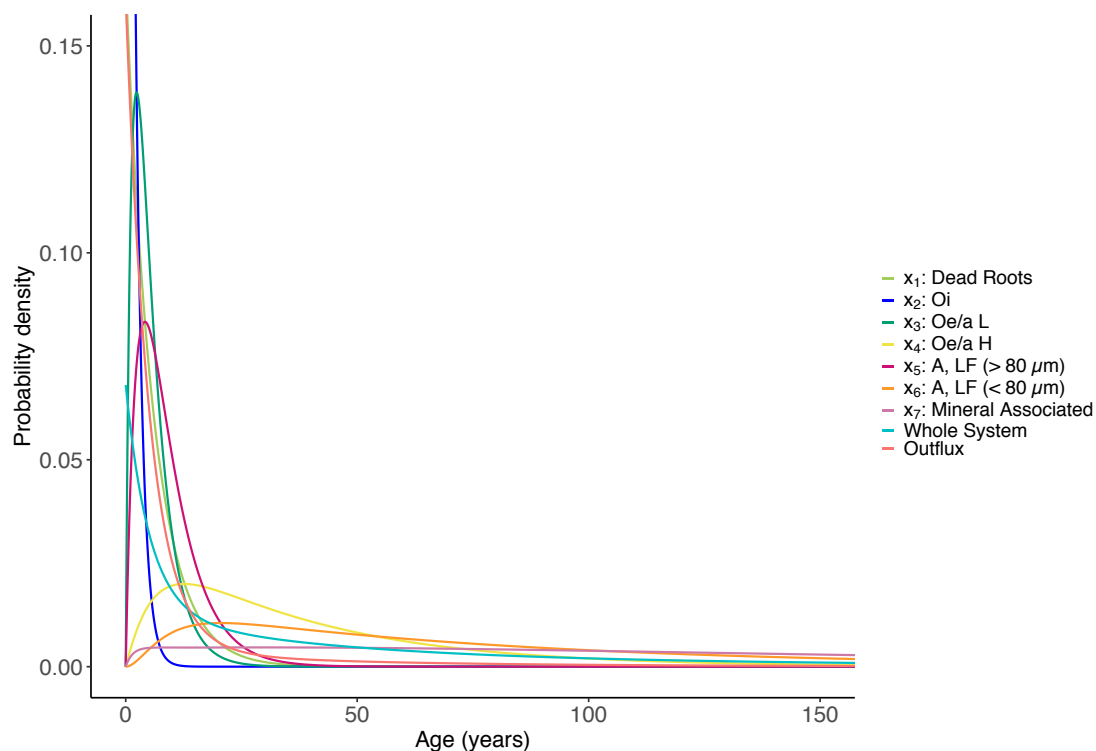
362 For the whole system in AD 1965 (Figure 7), the distribution of radiocarbon aggregates  
 363 the contributions of the different pools, which results in different peaks in the overall distri-  
 364 bution. The mode (i.e., the  $\Delta^{14}\text{C}$  with highest mass density) is below 0 ‰ because a large  
 365 portion of the total amount of carbon is contributed by the mineral associated pool that is  
 366 predominantly still pre-bomb carbon with little contribution from carbon fixed after 1964.  
 367 In addition, other pools that cycle fast, contribute relatively small amounts of bomb  $^{14}\text{C}$   
 368 to the overall distribution.

369 The radiocarbon distribution in the output flux in AD 1965 (Figure 7), i.e., the radio-  
 370 carbon distribution that corresponds to the transit time distribution for this year, has three  
 371 distinct peaks in the distribution. This distribution is very similar to that of the *Dead Roots*  
 372 pool (Figure S3), which is the main contributor to the total respiration flux. However, other  
 373 pools also contribute to the respiration flux with their radiocarbon signatures and emphasize  
 374 fluxes from the fastest cycling pool (*Oi*) and respiration of carbon that was present in other  
 375 pools before the bomb peak.

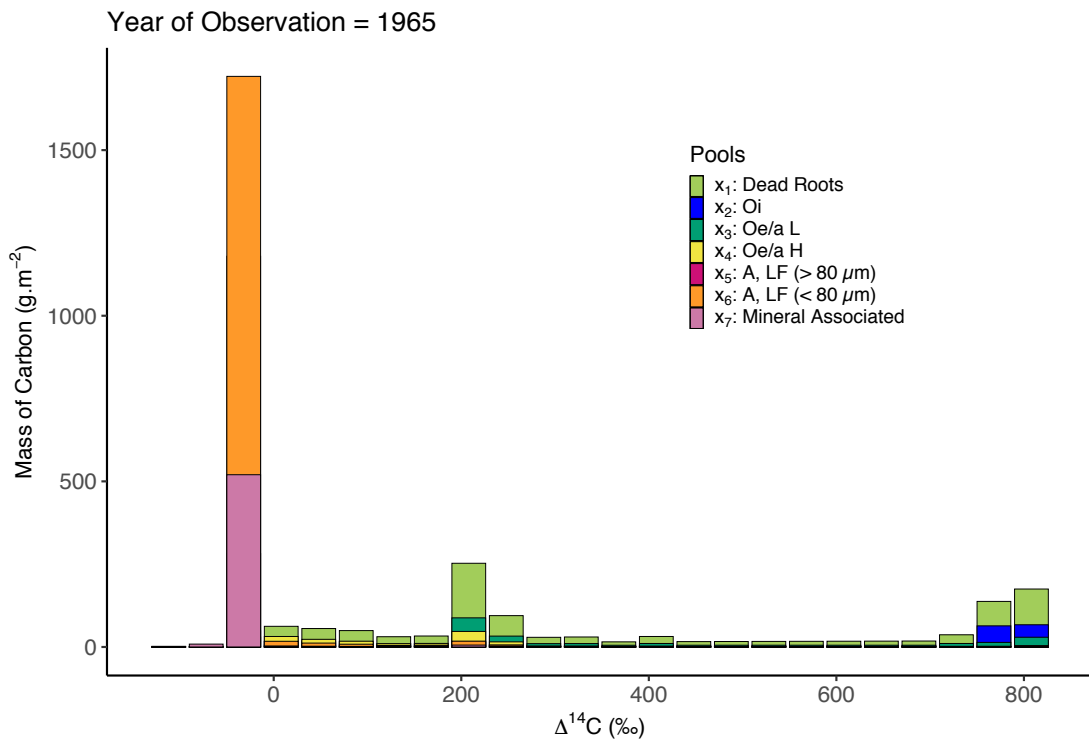
376 The shapes of the distributions change dramatically for subsequent years after the bomb  
 377 spike (Figure 5). For AD 2027, the expected  $\Delta^{14}\text{C}$  values of fast pools drop considerably, in  
 378 parallel with atmospheric  $^{14}\text{C}$ , compared to AD 1965. These fast pools do not stored much  
 379 radiocarbon from the bomb period, and their radiocarbon signatures reflect recent carbon  
 380 from the atmosphere. In contrast, slow cycling pools in AD 2027 had relatively high  $\Delta^{14}\text{C}$   
 381 values, mostly because they still contain radiocarbon from the bomb period. In the output  
 382 flux, as expected, since the respiration flux is dominated by the faster-cycling pools such

383 as *Dead Roots* and  $O_i$ , most of the radiocarbon is narrowly distributed around the recent  
 384 atmospheric  $\Delta^{14}\text{C}$  value in 2027, with almost no contributions from bomb  $^{14}\text{C}$ .

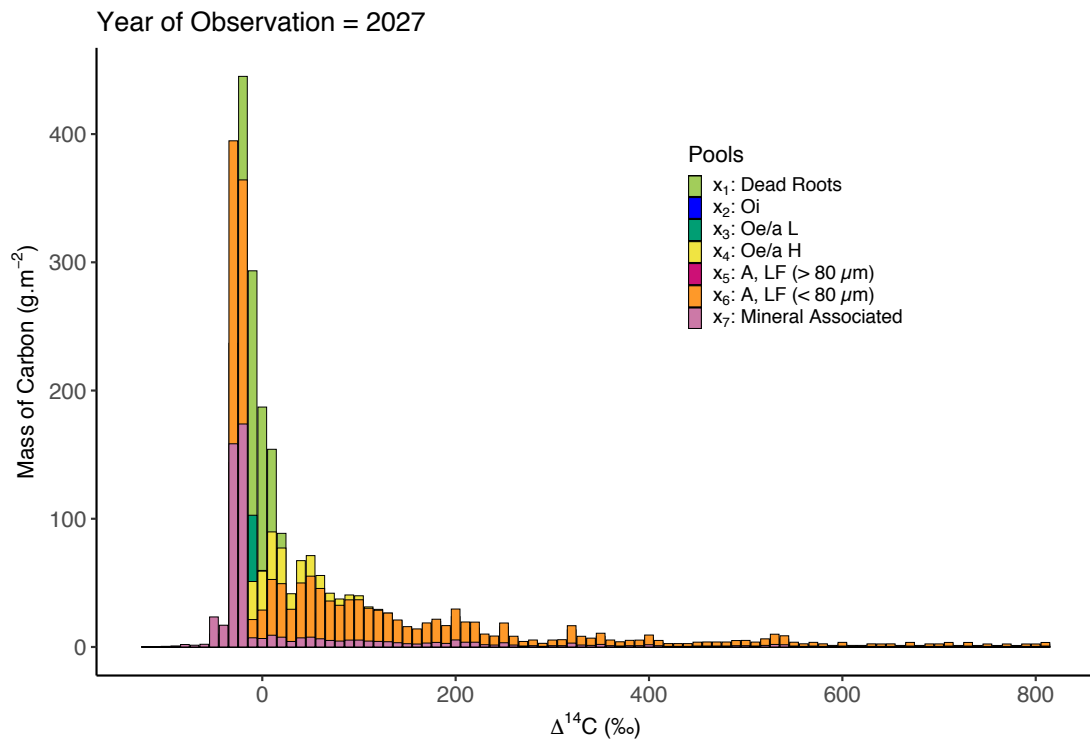
385 By the year 2100, the atmospheric  $\Delta^{14}\text{C}$  values have dropped to  $-254.5\text{‰}$  (Graven,  
 386 2015), reflecting the Suess effect. The distributions of most pools are less variable. Faster  
 387 cycling pools have dropped to reflect negative  $\Delta^{14}\text{C}$  in the atmosphere over the 73 years  
 388 since 2027, while the slow pools (*Mineral Associated*, *A*, *LF* ( $< 80\ \mu\text{m}$ ) and *Oe/a H* pools)  
 389 still show a wide range of  $\Delta^{14}\text{C}$  values that includes C fixed during the bomb period (now  
 390  $\sim 150$  years previously).



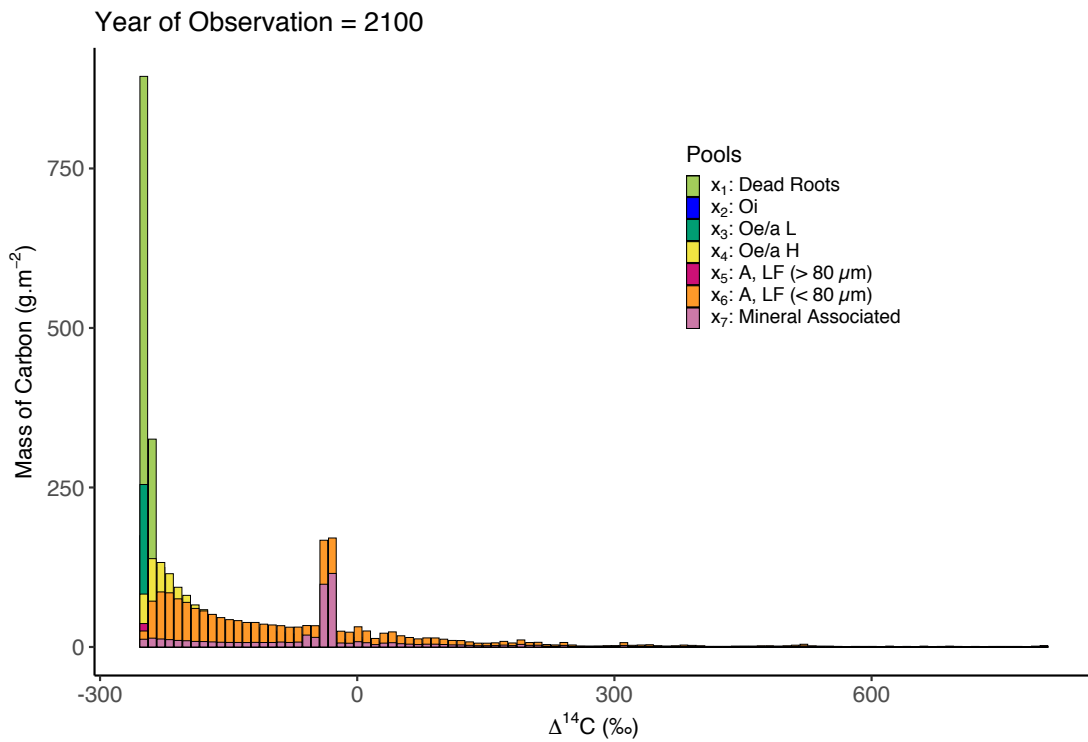
**Figure 3.** Age distributions for the Harvard Forest Soil model computed in a span of 1,000 years with a resolution of 0.1 year. The x-axis is limited to 150 years and the y-axis is limited to 0.15 for better visualization of the data.



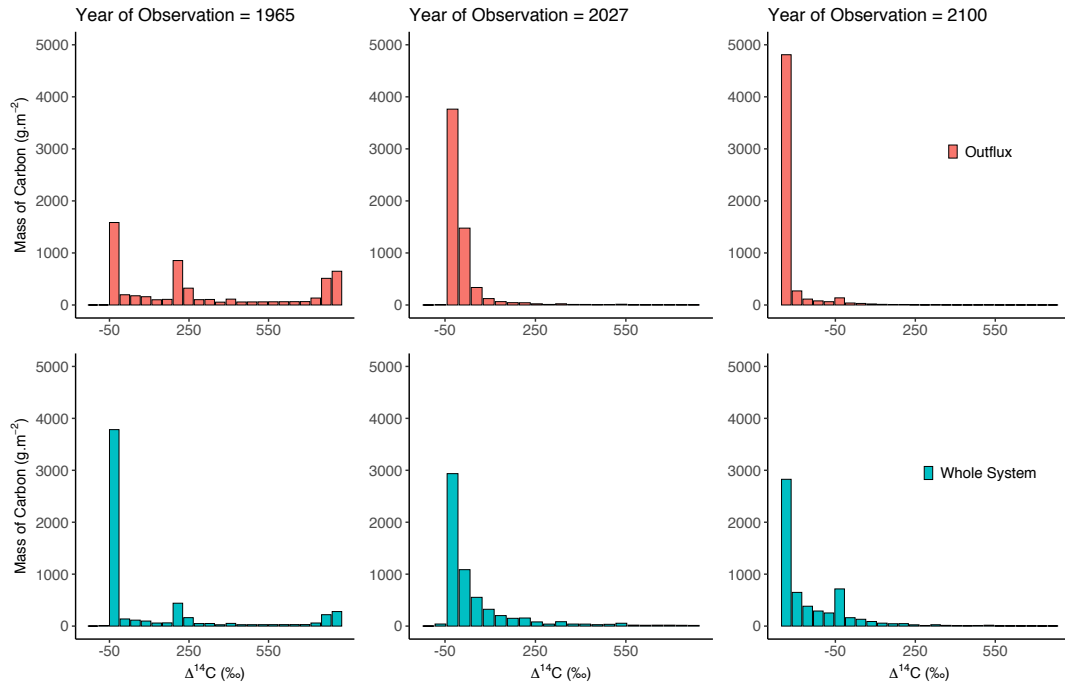
**Figure 4.**  $\Delta^{14}\text{C}$  distributions of each of the seven pools of the HFS model through the algorithm described above. The year of observation is 1965 – just after the bomb peak in 1964 – and the distributions are computed over 1,000 years. The bin size  $b$  is equal to 40 ‰. The expected value and standard deviation of this distribution is  $141 \pm 280$ .



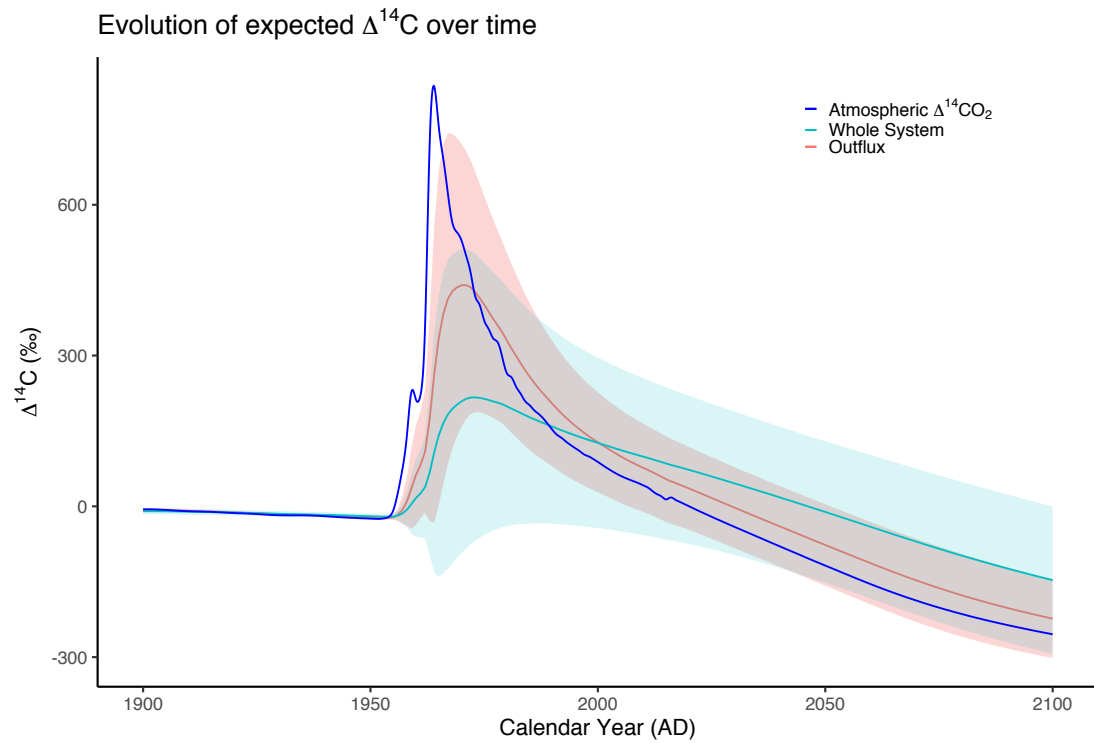
**Figure 5.**  $\Delta^{14}\text{C}$  distributions of each of the seven pools of the above-mentioned HFS model through the algorithm described above. The year of observation is 2027 and the distributions are computed over 1,000 years. The bin size  $b$  is equal to 10 ‰. The expected value and standard deviation of this distribution is  $54 \pm 144$ .



**Figure 6.**  $\Delta^{14}\text{C}$  distributions of each of the seven pools of the above-mentioned HFS model through the algorithm described above. The year of observation is 2100 and the distributions are computed over 1,000 years. The bin size  $b$  is equal to 10 ‰. The expected value and standard deviation of this distribution is  $-147 \pm 146$ .



**Figure 7.**  $\Delta^{14}\text{C}$  distributions of *Outflux* and *Whole System* of the HFS model for the years 1965, 2027 and 2100. The bin size  $b$  for all the three years is equal to 40 ‰.



**Figure 8.** Evolution of the expected  $\Delta^{14}\text{C}$  values of *Outflux* and *Whole System* for the HFS model between the years 1900 and 2100.



## 4.2 Comparison with measured data

Radiocarbon measurements of total soil CO<sub>2</sub> efflux at the Harvard Forest compared relatively well with the theoretical distributions of radiocarbon in the output flux obtained from our approach. Total soil CO<sub>2</sub> efflux includes both decomposition sources predicted by the model and root respiration, estimated by Gaudinski et al. (2000) to be ~ 55% and to have Δ<sup>14</sup>C values equal to the atmosphere in any given year.

For the years 1996, 1998, 2002, and 2008, the measurements were always within the expected range (Figure 9, Table 1). In all cases, the average of the measurements was relatively close to the expected value of the theoretical distributions. However, the variability of the observations was smaller than the expected variability from the model (Figure 10).

In particular, the expected values were systematically higher in <sup>14</sup>C than average of the observations for years 1996, 1998, 2002, and 2008 by 23.5 ‰, 21.8 ‰, 15.1 ‰, and 10.8 ‰, respectively (Figure 10).

There was one order of magnitude difference between the most prominent peak and secondary peaks of the theoretical Δ<sup>14</sup>C distributions. For the years 1996, 2002, and 2008, the main peak represents a mass of respired carbon within the year of ~ 10<sup>3</sup> g m<sup>-2</sup>. For the year 1998, the highest peak represents a mass of respired C of ~ 10<sup>2</sup> g m<sup>-2</sup>. For all those years, the bin size *b* was 10 ‰.

For the year 1996 (Figure 9a), the measurements of soil CO<sub>2</sub> efflux ranged from 104.3 ‰ to 167.3 ‰ (σ = 17.3 ‰). The theoretical most probable values also fall in this range: (112, 122]. Secondary peaks are comprised in ranges with magnitude of one bin size, starting in Δ<sup>14</sup>C values of -28 ‰ and 102 ‰. For this year, our theoretical estimations also show secondary peaks in a wide range of Δ<sup>14</sup>C values ranging from 122 ‰ to 212 ‰ (Table 1).

For the year 1998 (Figure 9b), the measurements of soil CO<sub>2</sub> efflux ranged from 66.4 ‰ to 193.9 ‰, (σ = 26.2 ‰). The main theoretical peaks are in Δ<sup>14</sup>C ranges below 0 ‰ – (-37, -17] – and above 0 ‰ – (93, 153]. The peak with negative Δ<sup>14</sup>C values does not appear in the soil CO<sub>2</sub> efflux measurements. Additionally, secondary peaks of theoretical estimations have a wide range of values, falling in the ranges: (153, 273], (323, 333], and (493, 503] in ‰ (Table 1).

For the year 2002 (Figure 9c), the measurements of soil CO<sub>2</sub> efflux range from 88 ‰ to 117.9 ‰, (σ = 8.4 ‰). Again, the theoretical prominent peak falls in the range: (82, 102]. One secondary peak includes the range observed in the empirical data – (102, 152]; however, additional secondary peaks of one bin size starting on -28 ‰ and 72 ‰ are not captured by the measurements (Table 1).

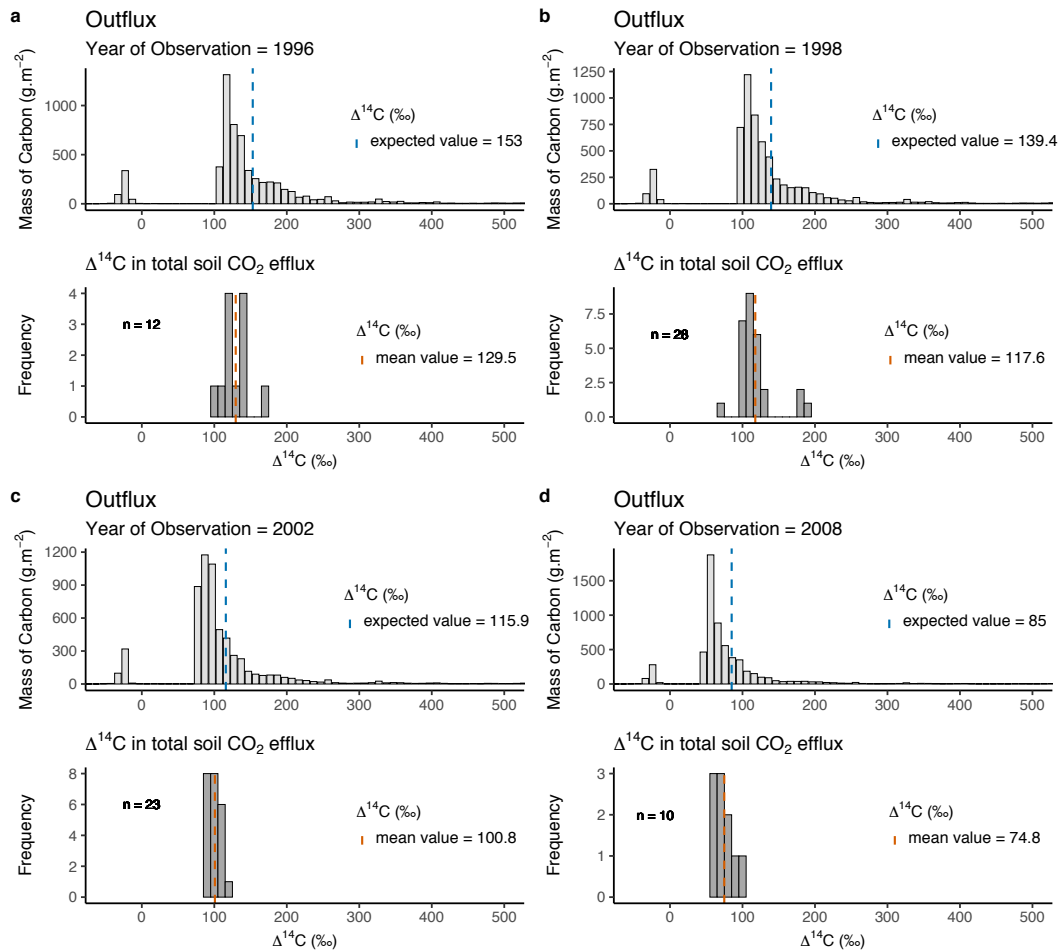
Finally, for the year 2008 (Figure 9d), the measurements of soil CO<sub>2</sub> efflux range from 60.8 ‰ to 104.7 ‰, (σ = 13.6 ‰). The peaks for this year are concentrated in the range (41, 121], with the highest density in the bin (51, 61] in ‰. An additional secondary peak falls in the negative part of the Δ<sup>14</sup>C axis, comprising values between -29 ‰ and -19 ‰ (Table 1).

The standard deviation of the observations were 17.3 ‰, 26.2 ‰, 8.4 ‰, and 13.6 ‰ for the years 1996, 1998, 2002, and 2008, respectively, which are smaller than the expected standard deviation of the distributions, which were 107.6 ‰, 103.3 ‰, 96.3 ‰, and 89.7 ‰ for the corresponding years.

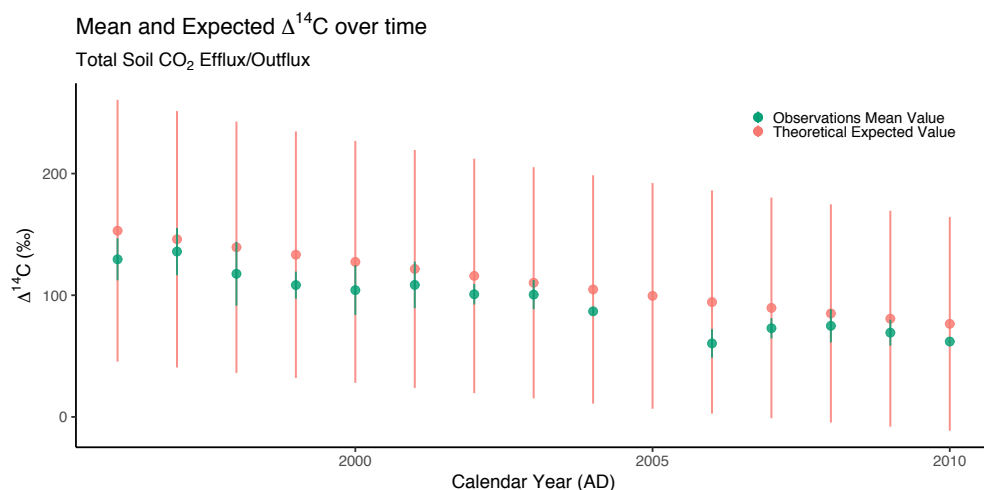
**Table 1.**  $\Delta^{14}\text{C}$  ranges with the highest masses of radiocarbon according to our estimations;  $\Delta^{14}\text{C}$  expected values according to weighted mean of mass distribution of radiocarbon; and observed  $\Delta^{14}\text{C}$  mean values of soil  $\text{CO}_2$  efflux.

$\Delta^{14}\text{C}$ [‰]				
Year	Primary Peaks <sup>a</sup>	Secondary Peaks <sup>b</sup>	Expected value <sup>c</sup>	Mean value <sup>d</sup>
1996	(112, 122]	(-28, -18], (102, 112], (122, 212]	$153 \pm 107.6$	$129.5 \pm 17.3$
1998	(-37, -17], (93, 153]	(153, 273], (323, 333], (493, 503]	$139.4 \pm 103.3$	$117.6 \pm 26.2$
2002	(82, 102]	(-28,-18], (72, 82], (102, 152]	$115.9 \pm 96.3$	$100.8 \pm 8.4$
2008	(51,61]	(41, 51], (61, 121], (-29, -19]	$85 \pm 89.7$	$74.8 \pm 13.6$

a For 1996, 2002 and 2008, masses  $\sim 10^3 \text{ g m}^{-2}$ ; For 1998, masses  $\sim 10^2 \text{ g m}^{-2}$ ;  
 b For 1996, 2002 and 2008, masses  $\sim 10^2 \text{ g m}^{-2}$ ; For 1998, masses  $\sim 10 \text{ g m}^{-2}$ ;  
 c Expected value of theoretical radiocarbon distribution of the Outflux (weighted mean);  
 d Mean value of the  $\Delta^{14}\text{C}$  values measured on soil  $\text{CO}_2$  efflux from the Harvard Forest.



**Figure 9.** Comparison between theoretical radiocarbon distribution and independent empirical data. **a:** Year of observation equals to AD 1996; **b:** Year of observation equals to AD 1998; **c:** Year of observation equals to AD 2002; **d:** Year of observation equals to AD 2008.



**Figure 10.** Change over time of the expected  $\Delta^{14}\text{C}$  values and of the mean  $\Delta^{14}\text{C}$  values of the total soil  $\text{CO}_2$  efflux in the Harvard Forest (HFS model).

434 **5 Discussion**

435 **5.1 How do distributions of radiocarbon change over time as a consequence**  
 436 **of changes in atmospheric radiocarbon?**

437 Our results clearly showed that distributions of radiocarbon in a compartmental system  
 438 change considerably over time, despite the stationarity of the age and transit time distribu-  
 439 tions for systems at steady-state, where the total mass of carbon does not change over time.  
 440 These changes reflect recent and expected dramatic changes in the isotopic signature of the  
 441 inputs originating in the atmosphere, including the bomb spike and fossil fuel dilution.

442 For fast cycling pools, we expect changes to match that of the radiocarbon content in  
 443 the atmosphere. Consequently, the radiocarbon distributions for fast cycling pools present  
 444 peaks in  $\Delta^{14}\text{C}$  values similar to those from the contemporary atmospheric radiocarbon. That  
 445 is an effect of the fast response of highly dynamic pools to the variations in the isotopic  
 446 composition of the system inputs. As fast pools are the major contributors to the output  
 447 flux, the total respiration also has similar narrow distributions close to the atmospheric  
 448  $\Delta^{14}\text{C}$  in the year of observation/sampling ( $t_0$ ).

449 For slow cycling pools that receive carbon from other pools, we expect wider distribu-  
 450 tions that include contributions from C fixed decades to centuries in the past. Thus, bomb  
 451  $^{14}\text{C}$  takes a longer time to be observed in the radiocarbon distributions.

452 As a consequence of fossil fuel ( $^{14}\text{C}$ -free) emissions to the atmosphere, dilution of at-  
 453 mospheric radiocarbon (*Suess effect*, Suess (1955)) is expected to affect future radiocarbon  
 454 distributions. This further widens distributions in slow cycling pools, and causes fast cy-  
 455 cling pools to have lower  $\Delta^{14}\text{C}$  values than slow cycling pools. The Suess effect become  
 456 particularly relevant in the distributions for future years, as shown in the distributions of  
 457 radiocarbon based on the forecast of atmospheric  $\Delta^{14}\text{C}$  values. The  $\Delta^{14}\text{C}$  in the atmo-  
 458 sphere is estimated to achieve values as low as ca.  $-254 \text{‰}$  in 2100 for the RCP8.5 scenario  
 459 (Graven, 2015). Such low values can appear with relatively high density in two cases: (i)  
 460 if the pool cycles fast but the  $\Delta^{14}\text{C}$  values in the atmosphere present high dilution (as in  
 461 2100), or (ii) with natural or bomb, however non-diluted,  $\Delta^{14}\text{C}$  values in the atmosphere,  
 462 but in very slow cycling pools (i.e.,  $>2,500$  yrs of carbon age). The latter case reflects  
 463 sufficient time for radioactive decay to reduce radiocarbon values in the carbon residing in

464 the system. In experiments, this could result in an inability to distinguish faster and slower  
465 cycling pools using  $\Delta^{14}\text{C}$  expected values. Thus one advantage of using these radiocarbon  
466 distributions is to get insight into the dynamics of transfers in the compartmental system,  
467 and to emphasize when these becomes less meaningful in the future years. Such issues can  
468 begin as soon as in 2027, when the  $\Delta^{14}\text{C}$  values start to decline to values never observed  
469 before by natural processes (i.e., without the anthropogenic effects such as the fossil fuel  
470 emissions). In the forecast for central Europe (Sierra, 2018), this transition year occurs as  
471 soon as 2022. This underlines the urgency of measurements in the current situation and  
472 the use of archived samples from the last decades, to emphasize the difference between fast  
473 pools that will track the changing atmosphere and slower pools that adjust more gradually  
474 and retain bomb  $^{14}\text{C}$  signals even in future decades.

## 475 **5.2 How do empirical data compare to these conceptual radiocarbon dis-** 476 **tributions?**

477 Measurements of radiocarbon in the output flux of a soil system, suggest that field  
478 measurements capture the mean value of the distributions, but not necessarily its variability.

479 Although we do not have independent observations available for specific pools to compare  
480 with our model predictions, we expect that for fast cycling pools the measurements  
481 will fall in a narrow range of  $\Delta^{14}\text{C}$  values, as can be observed in experiments assessing the  
482 fossil fuel  $\text{CO}_2$  distribution by measurements of  $\Delta^{14}\text{C}$  on deciduous leaves (Santos et al.,  
483 2019). For slow cycling pools, we would expect that the variability of  $\Delta^{14}\text{C}$  experimental  
484 data will be broader.

485 Carbon pools that cycle slowly can be very important for climate change mitigation,  
486 since they could store carbon for a longer time. Therefore, an accurate understanding of  
487 their dynamics is crucial. A valuable tool to assess these dynamics is using radiocarbon as  
488 a tracer to further constrain models. However, based on our results and interpretations,  
489 we believe that future research work should attempt at better capturing the variability of  
490 radiocarbon in such pools.

## 491 **5.3 What insights can these distributions provide for experimental and sam-** 492 **pling design for improving model-data comparisons by capturing the** 493 **entire variability of $\Delta^{14}\text{C}$ values?**

494 Overall, our results have implications for the interpretation of measured radiocarbon  
495 data and the design of empirical studies for improved understanding of carbon dynamics  
496 and comparison with models. The number of samples required to adequately represent the  
497 internal variability in radiocarbon depends on the year of observation and the particular  
498 compartment of interest. A priori determining sample sizes may be a suitable approach  
499 for future studies. For samples already collected, caution must be taken in interpreting  
500 the results, since a bulk measurement may not capture the whole distribution of possible  
501 radiocarbon values.

502 Our study opens up new opportunities to empirically determining radiocarbon distri-  
503 butions in compartmental systems. For example, this could be achieved by sampling designs  
504 that are representative of the compartments with higher variance, making sure the number  
505 of samples catches the entire potential variability. This way, it should be possible to  
506 determine empirical radiocarbon distributions.

507 Empirical determination of radiocarbon distributions in compartmental systems could  
508 be used to obtain age and transit time distributions using inverse statistical methods. This  
509 offers tremendous opportunities for accurate estimations of time metrics, incorporating the  
510 complexity of biological systems through multiple interconnected compartments. However,  
511 more research is still needed to determine whether radiocarbon distributions map to unique  
512 age and transit time distributions. To guarantee the uniqueness of the age and transit

513 time distributions from compartmental systems, one should be able to assure that only one  
514 combination of rates in the compartmental matrix builds the estimated distributions.

515 Moreover, as pointed out by Gaudinski et al. (2000), limited information about the  
516 cycling rates are obtained by  $^{14}\text{C}$  measurements of bulk SOM made at a single point in  
517 time. Therefore, being able to compute radiocarbon distributions for different years of  
518 observation could improve the interpretations of the time-evolution of soil radiocarbon in  
519 terms of carbon dynamics.

## 520 **6 Conclusion**

521 Compartmental models are a common approach to describe the dynamics of open sys-  
522 tems, particularly when modeling the carbon cycle in ecosystems. The mathematical equa-  
523 tions developed to obtain age and transit time distributions are a robust approach already  
524 used in several contexts and, therefore, using these distributions to obtain radiocarbon dis-  
525 tributions in the same systems is a powerful method. The algorithm presented, besides  
526 being simple, demonstrated the potential power of the method. It also showed how, for a  
527 specific model, predictions can be compared with experimental data.

528 Radiocarbon distributions can be used together with the known changes in atmospheric  
529  $\Delta^{14}\text{CO}_2$  to evaluate how models predict the changing distributions of radiocarbon in each  
530 compartment and its output over that last decades. This provides a powerful method to  
531 test models against observations and to refine model representations of C dynamics in soils  
532 and ecosystems.

533 Our results also have shown that the heterogeneity of the ecosystems described through  
534 the mixing of matter in the pools, is related to the shapes of the radiocarbon distributions.  
535 As opposed to age and transit time distributions for systems in steady-state, radiocarbon  
536 distributions are expected to vary over time, strongly depending on the year of observation  
537 as a consequence of the dependence on the atmospheric  $^{14}\text{C}$  input in the system. Thus,  
538 not only the distributions' shapes will change according to the year of observation, but also  
539 their expected values, modes, and variance.

540 Overall, fast cycling pools with less heterogeneity present narrow shapes for all the years  
541 of observations, whereas slower cycling pools as well as more heterogeneous compartments  
542 present wider shapes and multiple peaks of  $\Delta^{14}\text{C}$  for high labelled years (e.g., 1965, when  
543 the concentrations of  $^{14}\text{C}$  in the atmosphere were almost two times higher than the natural  
544 levels).

545 The theoretical distributions can be estimated for specific time points, however, that is  
546 not always feasible in experiments. That means the estimations through the algorithm have  
547 to be taken carefully when one aims to compare them to empirical data. It is also important  
548 to be aware of the radiocarbon atmospheric values used to estimate the distributions, as the  
549 variation of atmospheric  $\Delta^{14}\text{C}$  can influence the shapes and mean values of the distributions.  
550 In this sense, having accurate data on the  $^{14}\text{C}$  contents in the atmosphere is key for the  
551 determination of the radiocarbon distributions in multiple interconnected compartmental  
552 systems.

## 553 **Acknowledgments**

554 The authors would like to thank Ingeborg Levin for the meaningful comments and sug-  
555 gestions on this work. This work is also stimulated by the scientific research developed at  
556 the Amazon Tall Tower Observatory (ATTO), which is partly funded by the German Fed-  
557 eral Ministry of Education and Research (grant number 01LK1602A) and the Max Planck  
558 Society.

559 The atmospheric  $\Delta^{14}\text{CO}_2$  datasets used in this research are available through Graven  
 560 (2015), Graven et al. (2017), and Reimer et al. (2020). Data on the compartmental model  
 561 presented in this research, including the independent  $\Delta^{14}\text{C}$  data used for comparisons with  
 562 our estimations are available through Sierra, Trumbore, et al. (2012).

## 563 References

- 564 Bolin, B., & Rodhe, H. (1973). A note on the concepts of age distribution and transit time  
 565 in natural reservoirs. *Tellus*, *25*(1), 58–62.
- 566 Gaudinski, J. B., Trumbore, S. E., Davidson, E. A., & Zheng, S. (2000). Soil carbon cycling  
 567 in a temperate forest: radiocarbon-based estimates of residence times, sequestration  
 568 rates and partitioning of fluxes. *Biogeochemistry*, *51*(1), 33–69.
- 569 Goudriaan, J. (1992). Biosphere structure, carbon sequestering potential and the atmo-  
 570 spheric 14c carbon record. *Journal of Experimental Botany*, *43*(8), 1111–1119.
- 571 Graven, H. D. (2015). Impact of fossil fuel emissions on atmospheric radiocarbon and various  
 572 applications of radiocarbon over this century. *Proceedings of the National Academy of*  
 573 *Sciences*, *112*(31), 9542–9545.
- 574 Graven, H. D., Allison, C. E., Etheridge, D. M., Hammer, S., Keeling, R. F., Levin, I., ...  
 575 others (2017). Compiled records of carbon isotopes in atmospheric co2 for historical  
 576 simulations in cmip6. *Geoscientific Model Development (Online)*, *10*(12).
- 577 Hogg, A. G., Heaton, T. J., Hua, Q., Palmer, J. G., Turney, C. S., Southon, J., ... others  
 578 (2020). Shcal20 southern hemisphere calibration, 0–55,000 years cal bp. *Radiocarbon*,  
 579 *62*(4), 759–778.
- 580 Hogg, A. G., Hua, Q., Blackwell, P. G., Niu, M., Buck, C. E., Guilderson, T. P., ... others  
 581 (2013). Shcal13 southern hemisphere calibration, 0–50,000 years cal bp. *Radiocarbon*,  
 582 *55*(4), 1889–1903.
- 583 Hua, Q., Barbetti, M., & Rakowski, A. Z. (2013). Atmospheric radiocarbon for the period  
 584 1950–2010. *Radiocarbon*, *55*(4), 2059–2072.
- 585 Jacquez, J. A., & Simon, C. P. (1993). Qualitative theory of compartmental systems. *Siam*  
 586 *Review*, *35*(1), 43–79.
- 587 Jain, A. K., Kheshgi, H. S., & Wuebbles, D. J. (1997). Is there an imbalance in the global  
 588 budget of bomb-produced radiocarbon? *Journal of Geophysical Research: Atmo-*  
 589 *spheres*, *102*(D1), 1327–1333.
- 590 Lawrence, C. R., Beem-Miller, J., Hoyt, A. M., Monroe, G., Sierra, C. A., Heckman, K.,  
 591 ... others (2020). An open source database for the synthesis of soil radiocarbon data:  
 592 Israd version 1.0. *Earth System Science Data*, *12*(LLNL-JRNL-775042).
- 593 Levin, I., & Kromer, B. (1997). Twenty years of atmospheric 14 co 2 observations at  
 594 schauinsland station, germany. *Radiocarbon*, *39*(2), 205–218.
- 595 Levin, I., Münnich, K., & Weiss, W. (1980). The effect of anthropogenic co 2 and 14 c  
 596 sources on the distribution of 14 c in the atmosphere. *Radiocarbon*, *22*(2), 379–391.
- 597 Levin, I., Naegler, T., Kromer, B., Diehl, M., Francey, R., Gomez-Pelaez, A., ... Worthy,  
 598 D. (2010). Observations and modelling of the global distribution and long-term trend  
 599 of atmospheric 14co2. *Tellus B: Chemical and Physical Meteorology*, *62*(1), 26–46.
- 600 Metzler, H., & Sierra, C. A. (2017, jul). Linear autonomous compartmental models as  
 601 continuous-time markov chains: Transit-time and age distributions. *Mathematical*  
 602 *Geosciences*, *50*(1), 1–34. Retrieved from <https://doi.org/10.1007/s11004-017-9690-1>  
 603 doi: 10.1007/s11004-017-9690-1
- 604 Naegler, T., Ciais, P., Rodgers, K., & Levin, I. (2006). Excess radiocarbon constraints  
 605 on air-sea gas exchange and the uptake of co2 by the oceans. *Geophysical Research*  
 606 *Letters*, *33*(11).
- 607 Randerson, J., Enting, I., Schuur, E., Caldeira, K., & Fung, I. (2002). Seasonal and  
 608 latitudinal variability of troposphere  $\delta^{14}\text{CO}_2$ : Post bomb contributions from fossil fuels,  
 609 oceans, the stratosphere, and the terrestrial biosphere. *Global Biogeochemical Cycles*,  
 610 *16*(4).

611 Rasmussen, M., Hastings, A., Smith, M. J., Agosto, F. B., Chen-Charpentier, B. M., Hoff-  
 612 man, F. M., ... others (2016). Transit times and mean ages for nonautonomous  
 613 and autonomous compartmental systems. *Journal of mathematical biology*, *73*(6-7),  
 614 1379–1398.

615 Reimer, P. J., Austin, W. E., Bard, E., Bayliss, A., Blackwell, P. G., Ramsey, C. B., ...  
 616 others (2020). The intcal20 northern hemisphere radiocarbon age calibration curve  
 617 (0–55 cal kbp). *Radiocarbon*, *62*(4), 725–757.

618 Reimer, P. J., Bard, E., Bayliss, A., Beck, J. W., Blackwell, P. G., Ramsey, C. B., ... others  
 619 (2013). Intcal13 and marine13 radiocarbon age calibration curves 0–50,000 years cal  
 620 bp. *Radiocarbon*, *55*(4), 1869–1887.

621 Santos, G. M., Oliveira, F. M., Park, J., Sena, A. C., Chiquetto, J. B., Macario, K. D.,  
 622 & Grainger, C. S. (2019). Assessment of the regional fossil fuel co2 distribution  
 623 through  $\delta^{14}c$  patterns in ipê leaves: The case of rio de janeiro state, brazil. *City and*  
 624 *Environment Interactions*, *1*, 100001.

625 Sierra, C. (2018). Forecasting atmospheric radiocarbon decline to pre-bomb values. *Radio-*  
 626 *carbon*, *60*(4), 1055–1066.

627 Sierra, C., Ceballos-Núñez, V., Metzler, H., & Müller, M. (2018). Representing and un-  
 628 derstanding the carbon cycle using the theory of compartmental dynamical systems.  
 629 *Journal of Advances in Modeling Earth Systems*, *10*(8), 1729–1734.

630 Sierra, C., Müller, M., Metzler, H., Manzoni, S., & Trumbore, S. (2017). The muddle of  
 631 ages, turnover, transit, and residence times in the carbon cycle. *Global change biology*,  
 632 *23*(5), 1763–1773.

633 Sierra, C., Müller, M., & Trumbore, S. (2014). Modeling radiocarbon dynamics in soils:  
 634 Soilr version 1.1. *Geoscientific Model Development*, *7*(5), 1919–1931.

635 Sierra, C., Müller, M., Trumbore, S., et al. (2012). Models of soil organic matter decomposi-  
 636 tion: the soilr package, version 1.0. *Geoscientific Model Development*, *5*, 1045–1060.

637 Sierra, C., Trumbore, S., Davidson, E., Frey, S., Savage, K., & Hopkins, F. (2012). Pre-  
 638 dicting decadal trends and transient responses of radiocarbon storage and fluxes in a  
 639 temperate forest soil. *Biogeosciences*, *9*(8), 3013–3028.

640 Stuiver, M., & Polach, H. A. (1977). Discussion reporting of  $^{14}c$  data. *Radiocarbon*, *19*(3),  
 641 355–363.

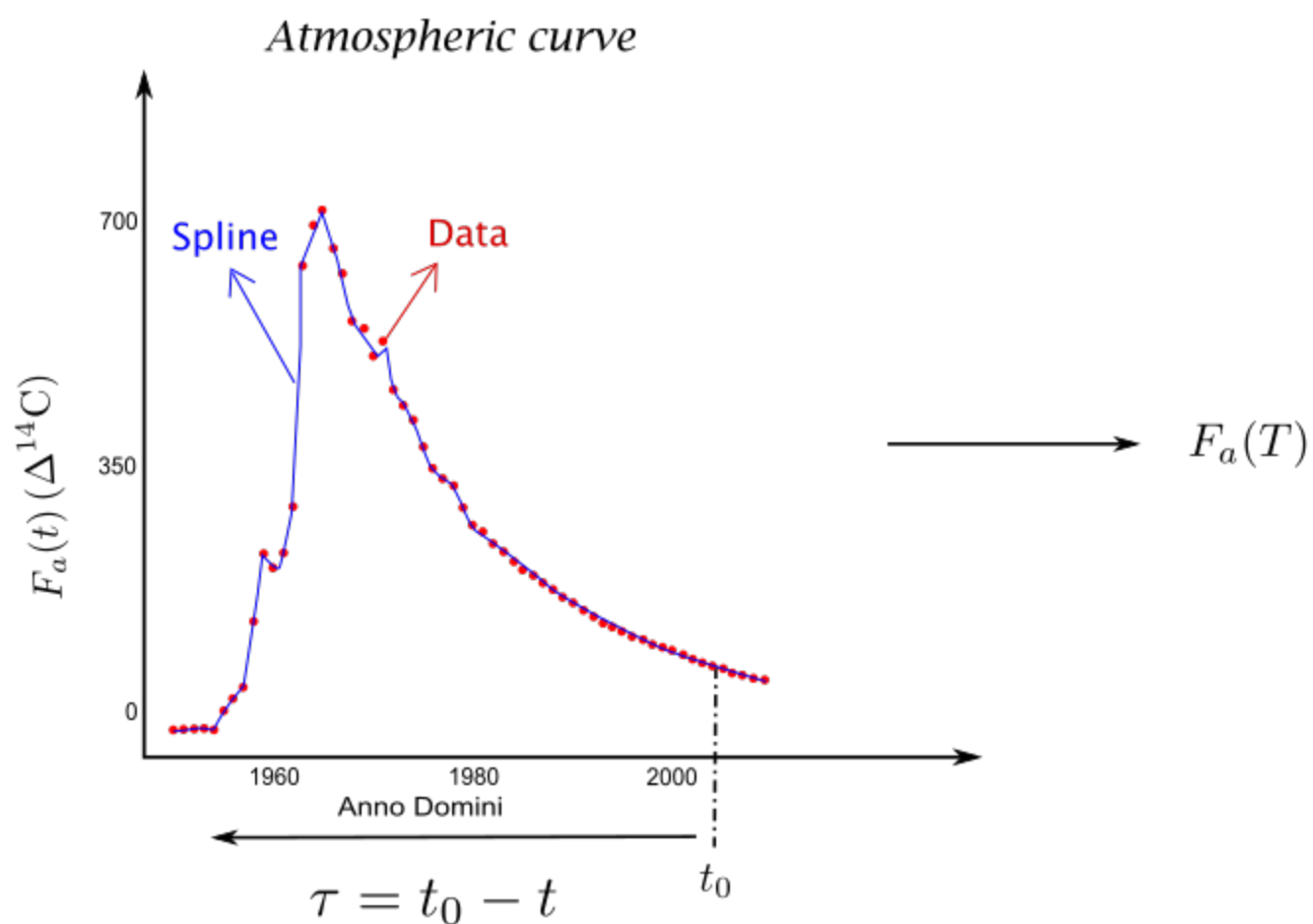
642 Suess, H. E. (1955). Radiocarbon concentration in modern wood. *Science*, *122*(3166),  
 643 415–417.

644 Thompson, M. V., & Randerson, J. T. (1999). Impulse response functions of terrestrial  
 645 carbon cycle models: method and application. *Global Change Biology*, *5*(4), 371–  
 646 394.

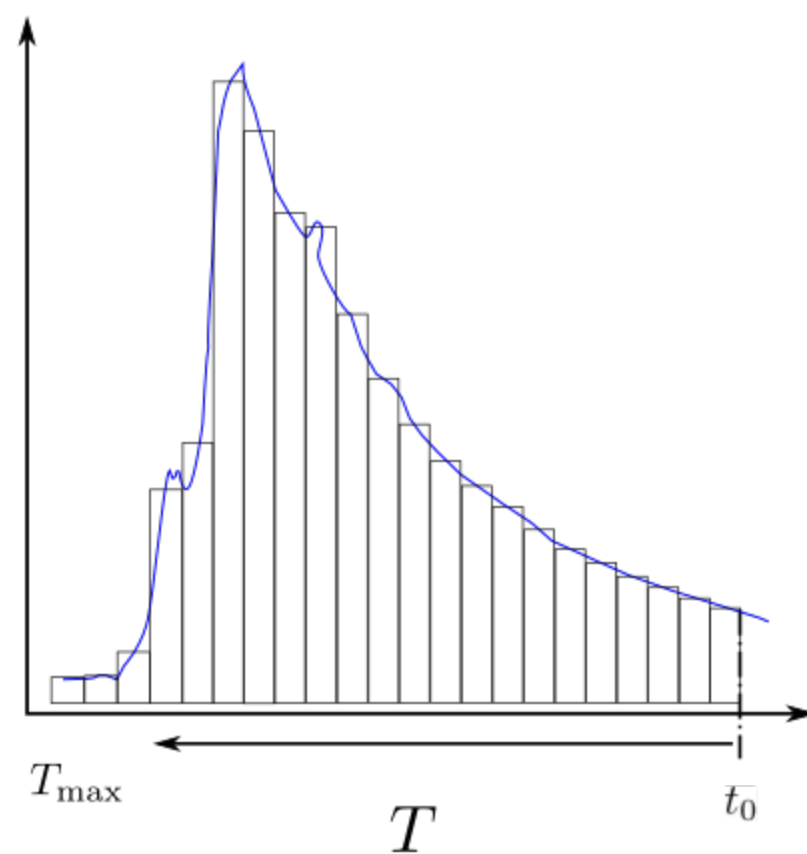
**Figure 1.**



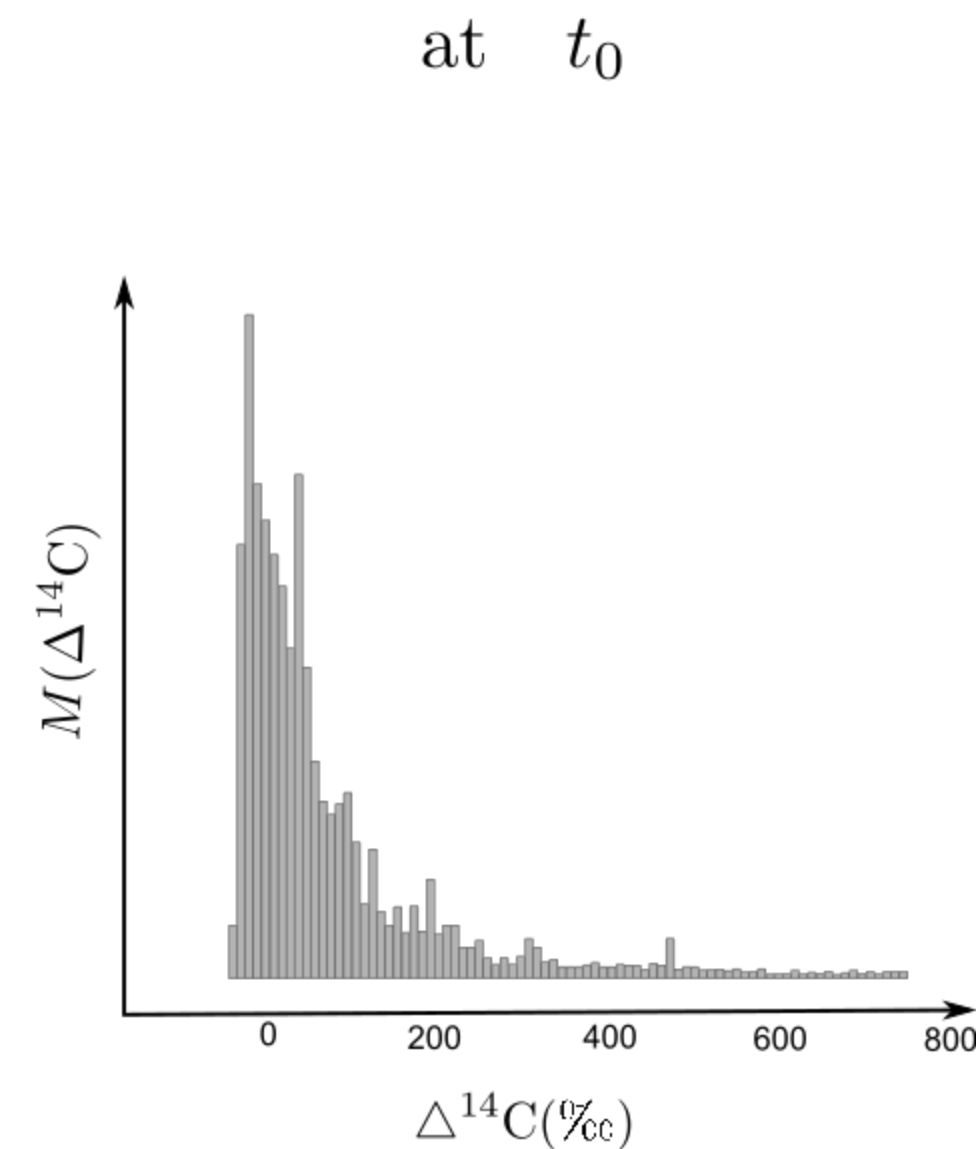
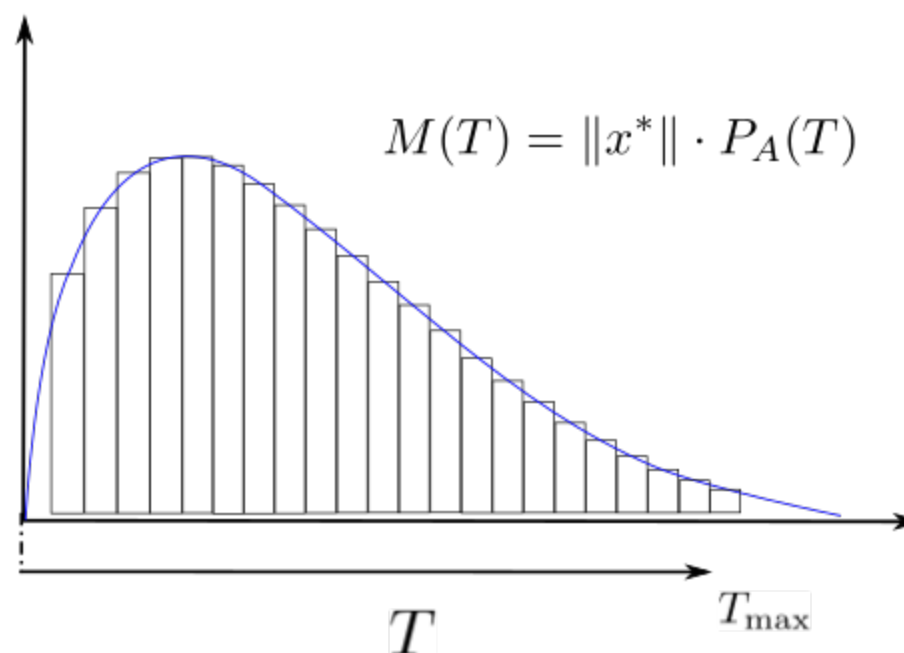
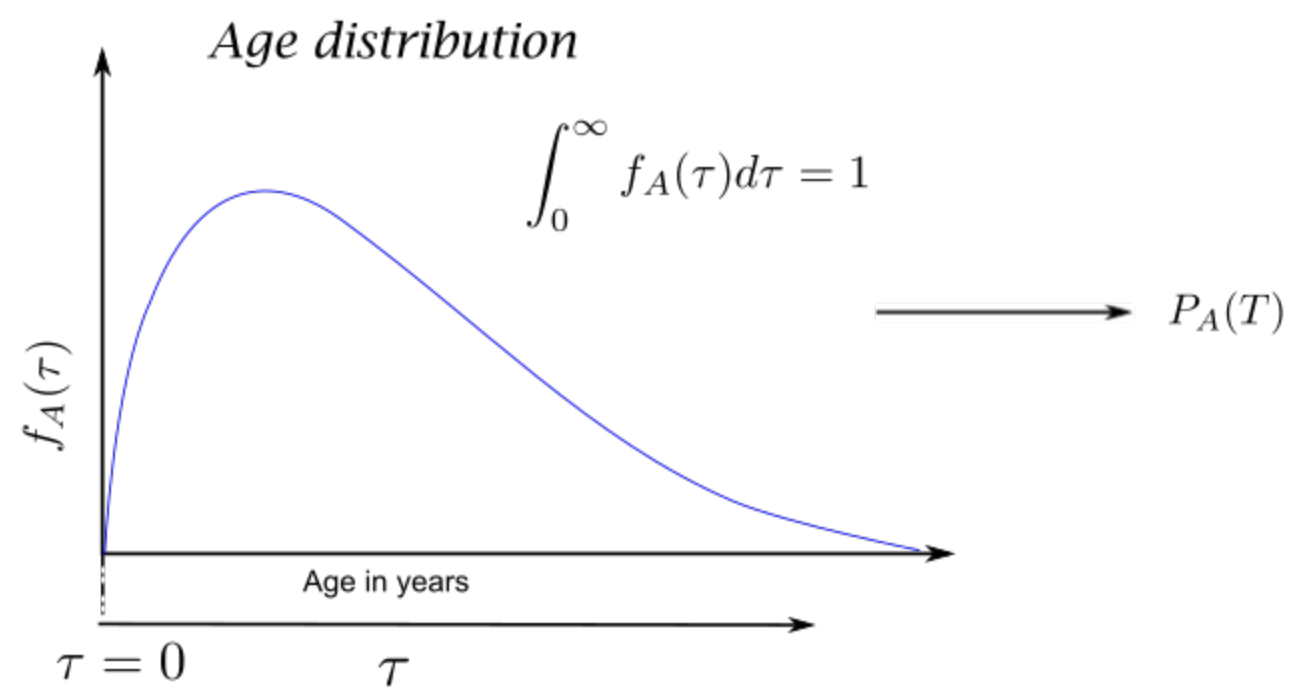
# 1. Homogenization



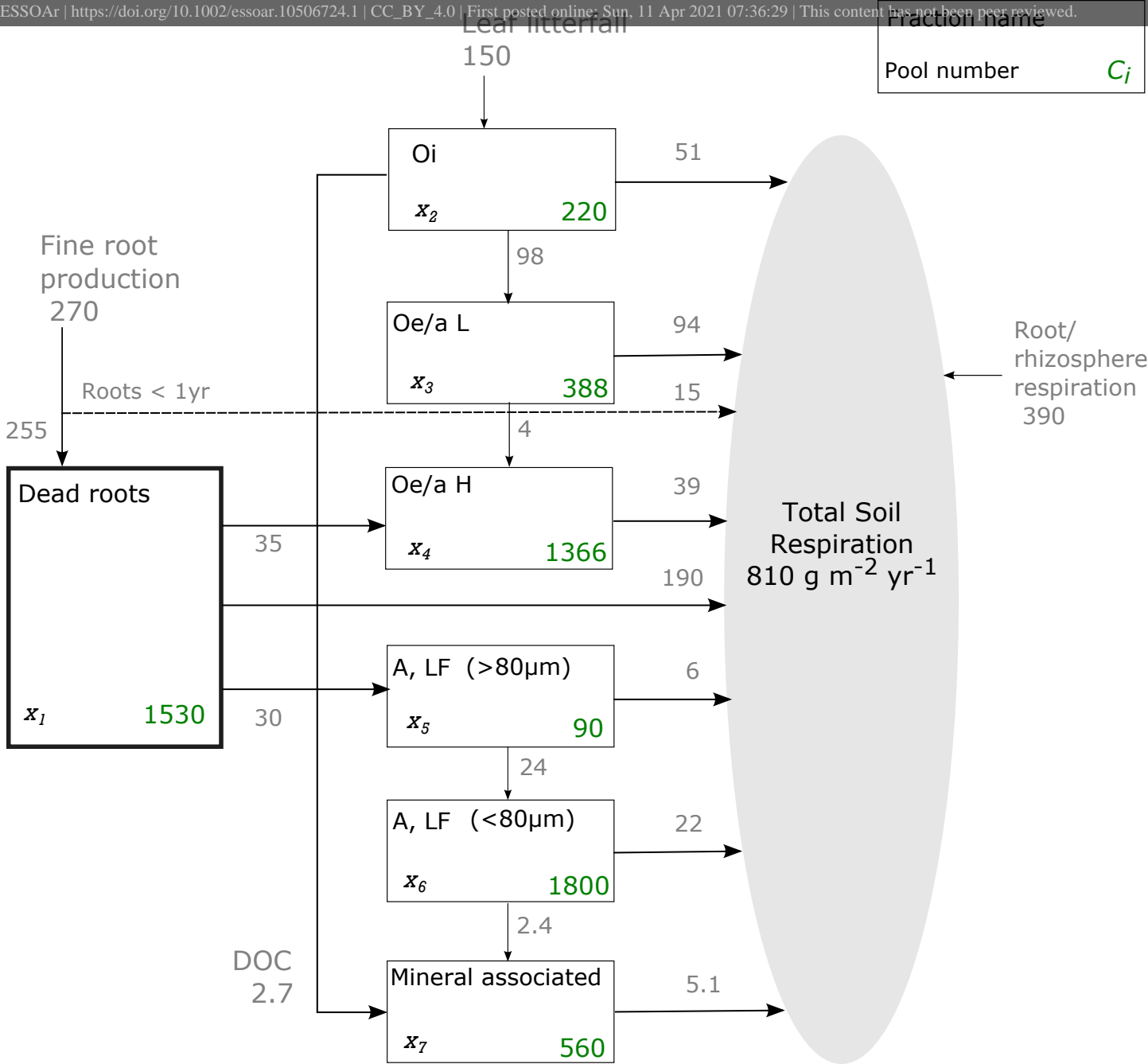
# 2. Discretization



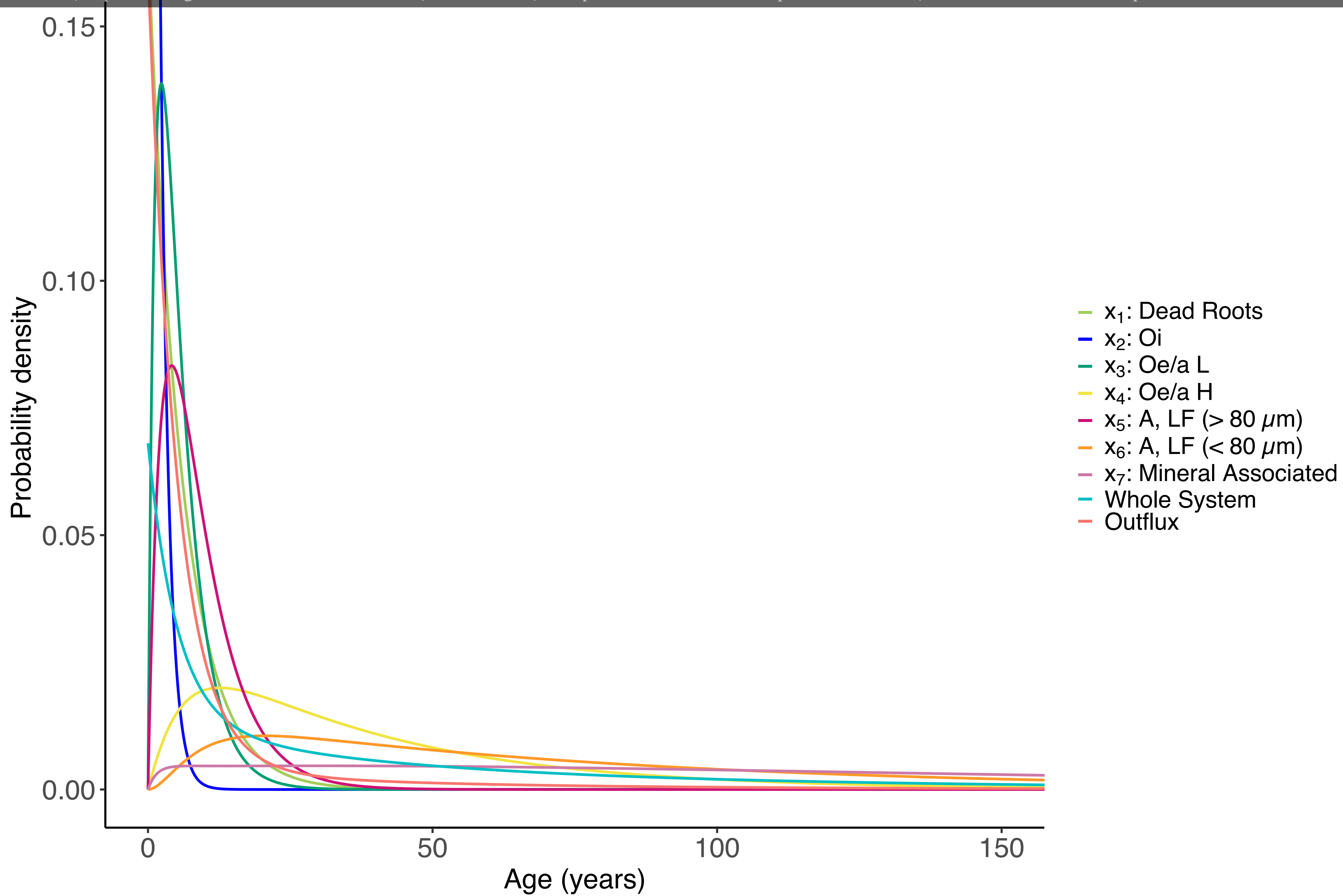
# 3. Aggregation



**Figure 2.**

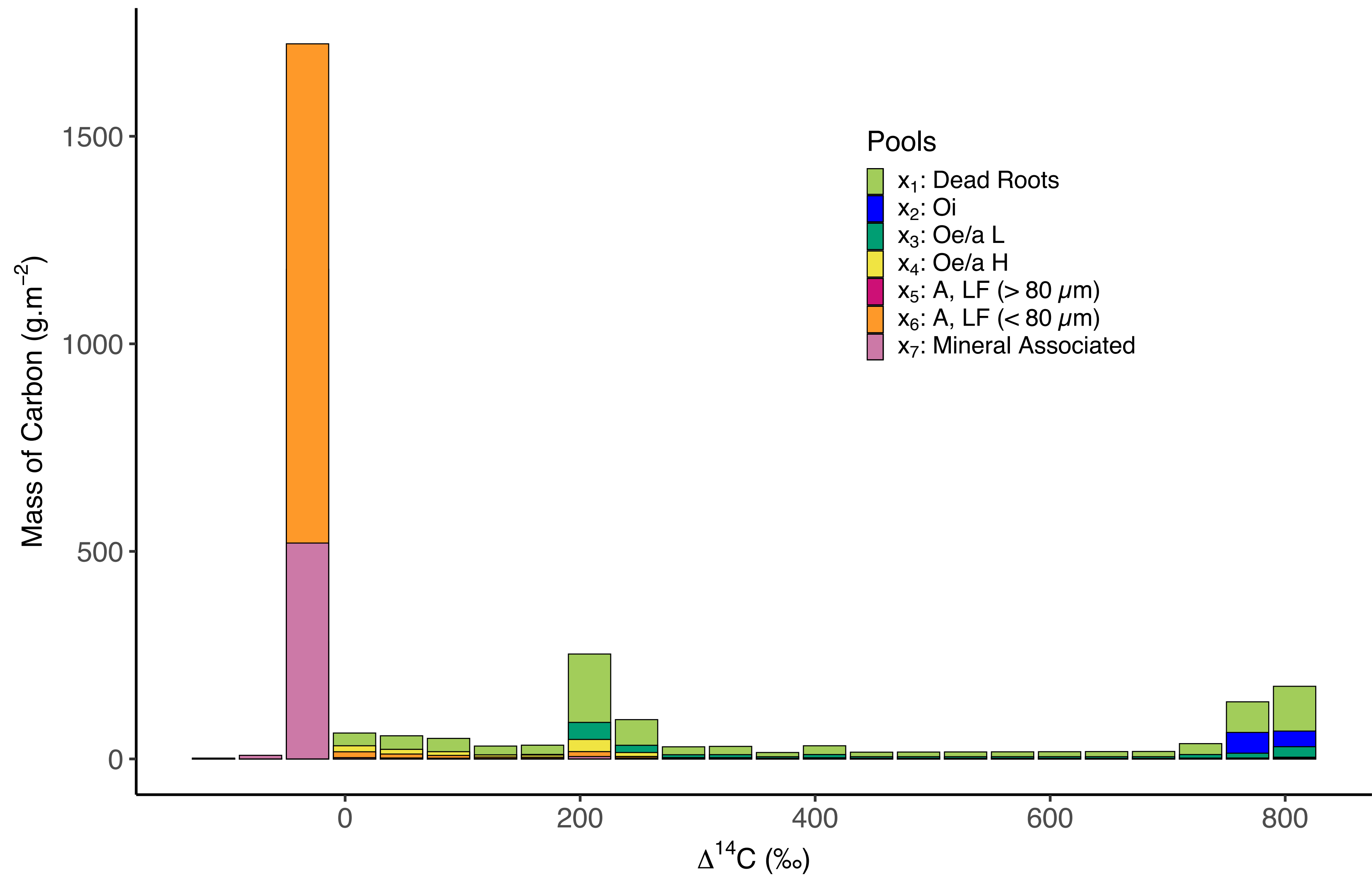


**Figure 3.**



**Figure 4.**

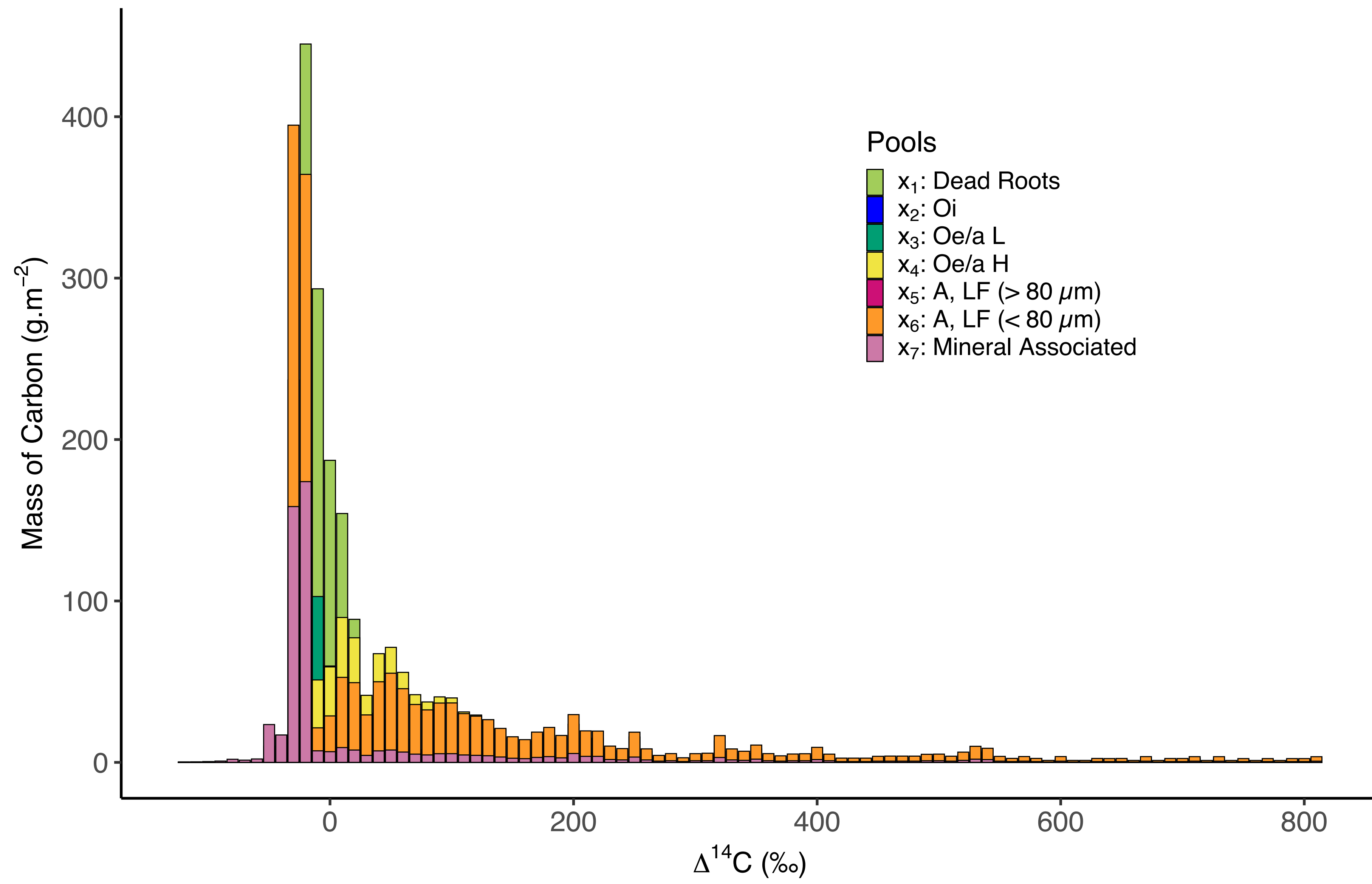
Year of Observation = 1965



**Figure 5.**

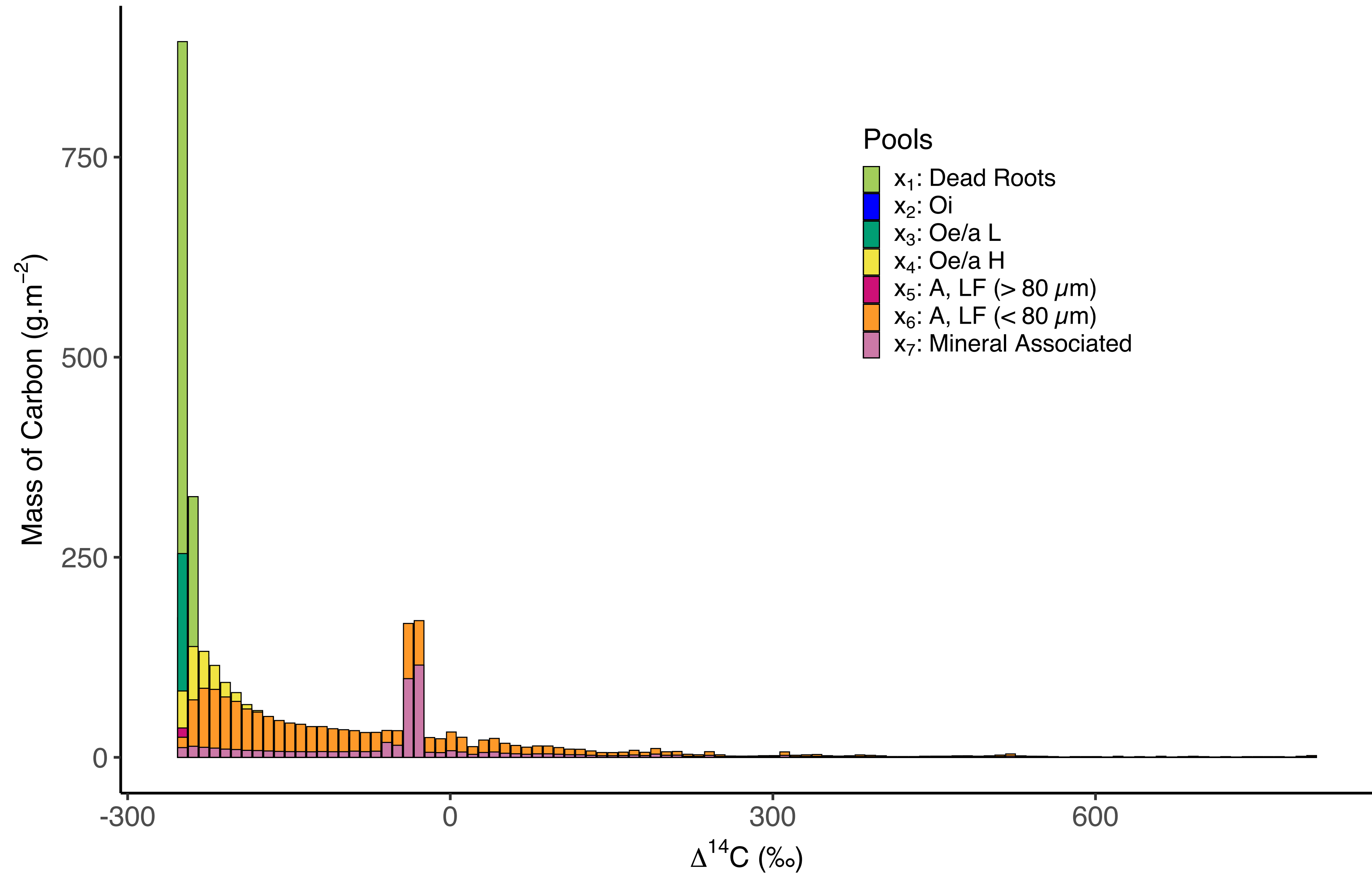


Year of Observation = 2027

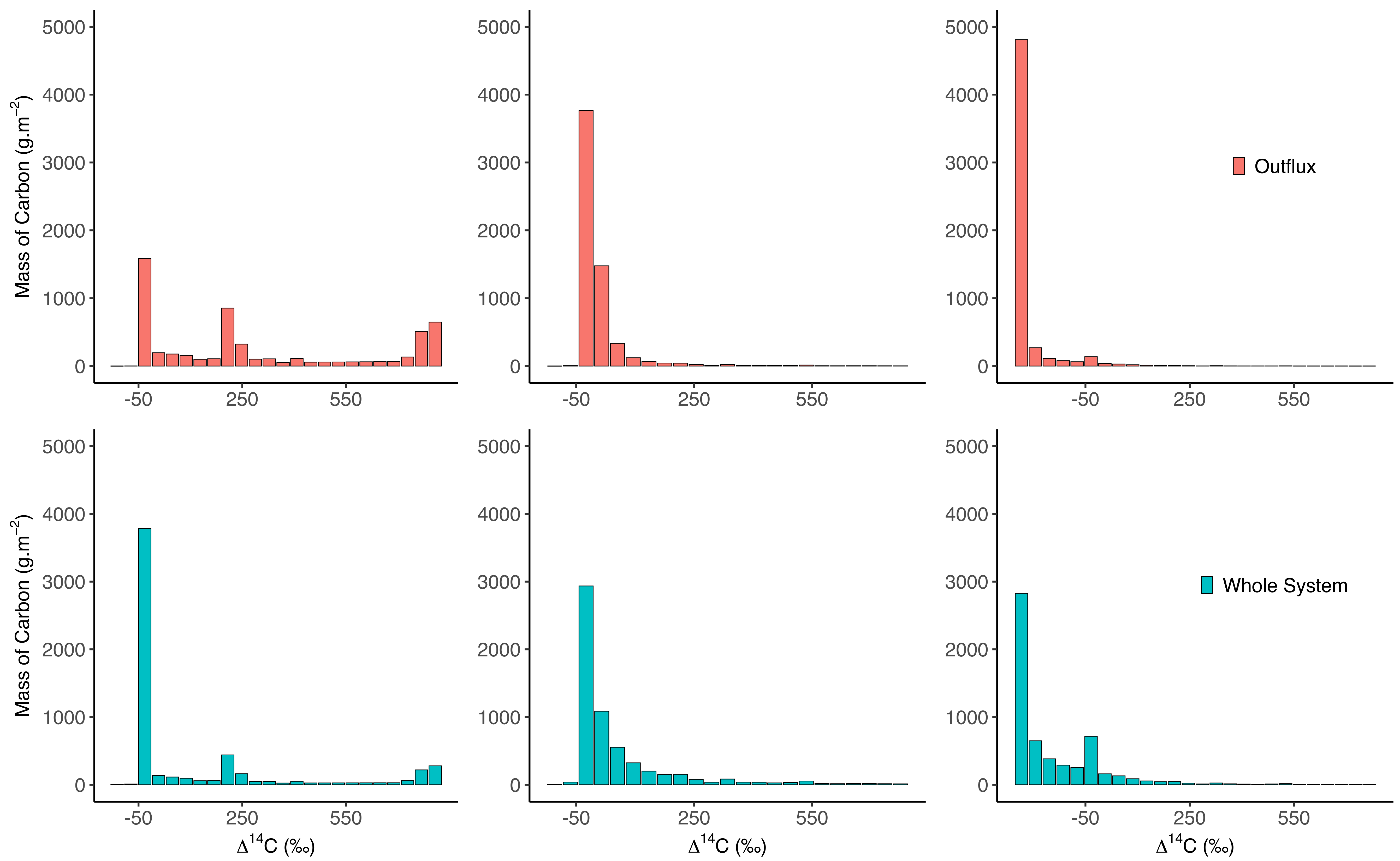


**Figure 6.**

Year of Observation = 2100

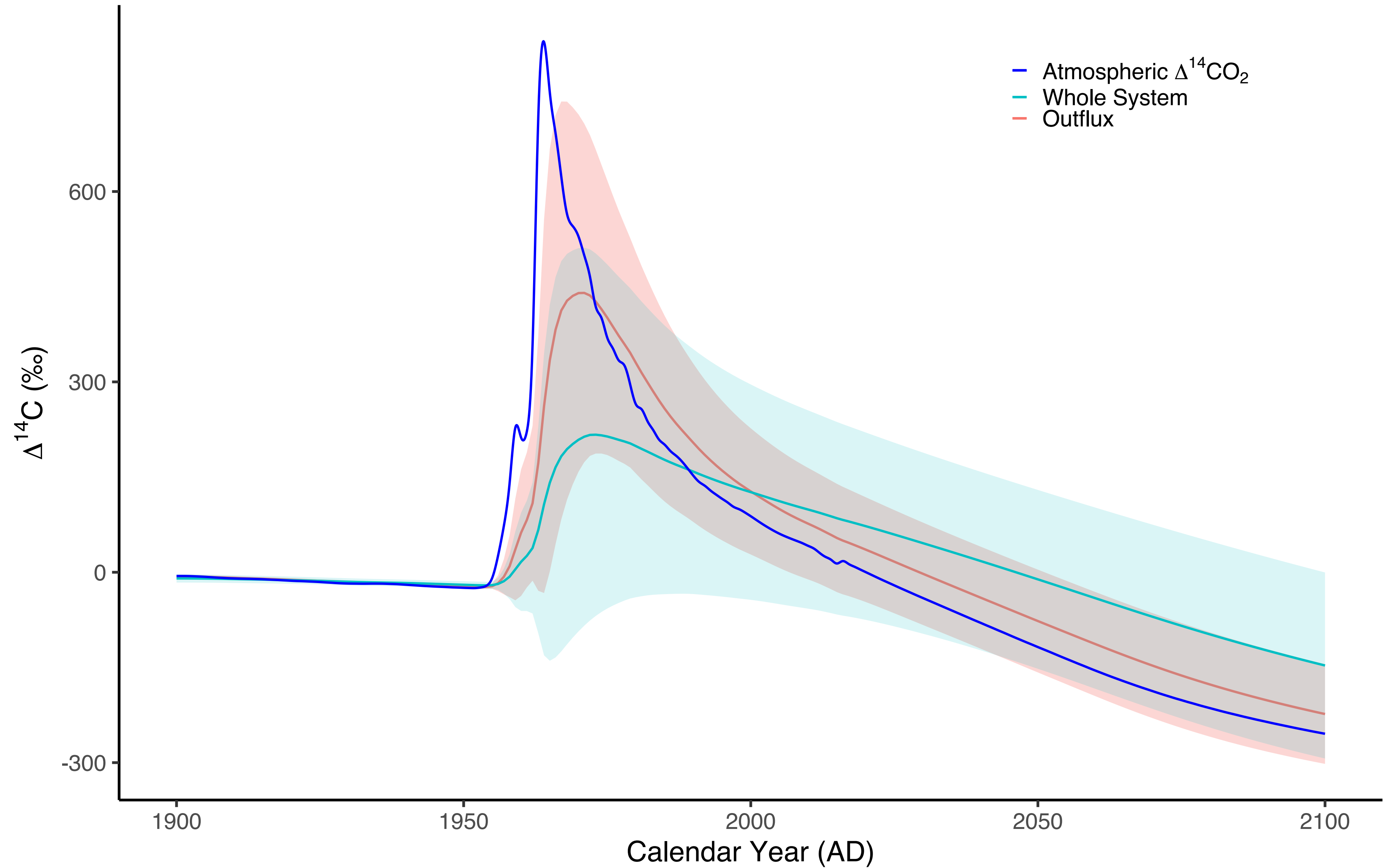


**Figure 7.**



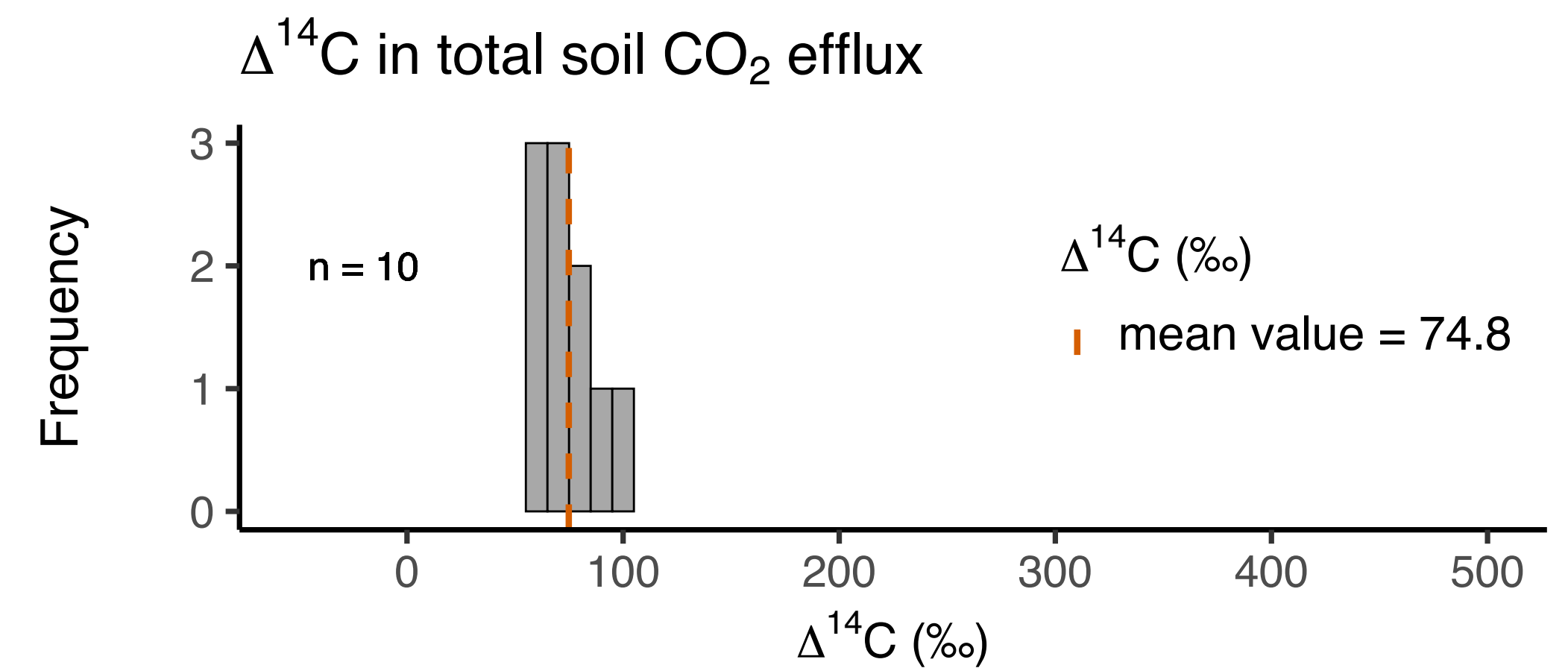
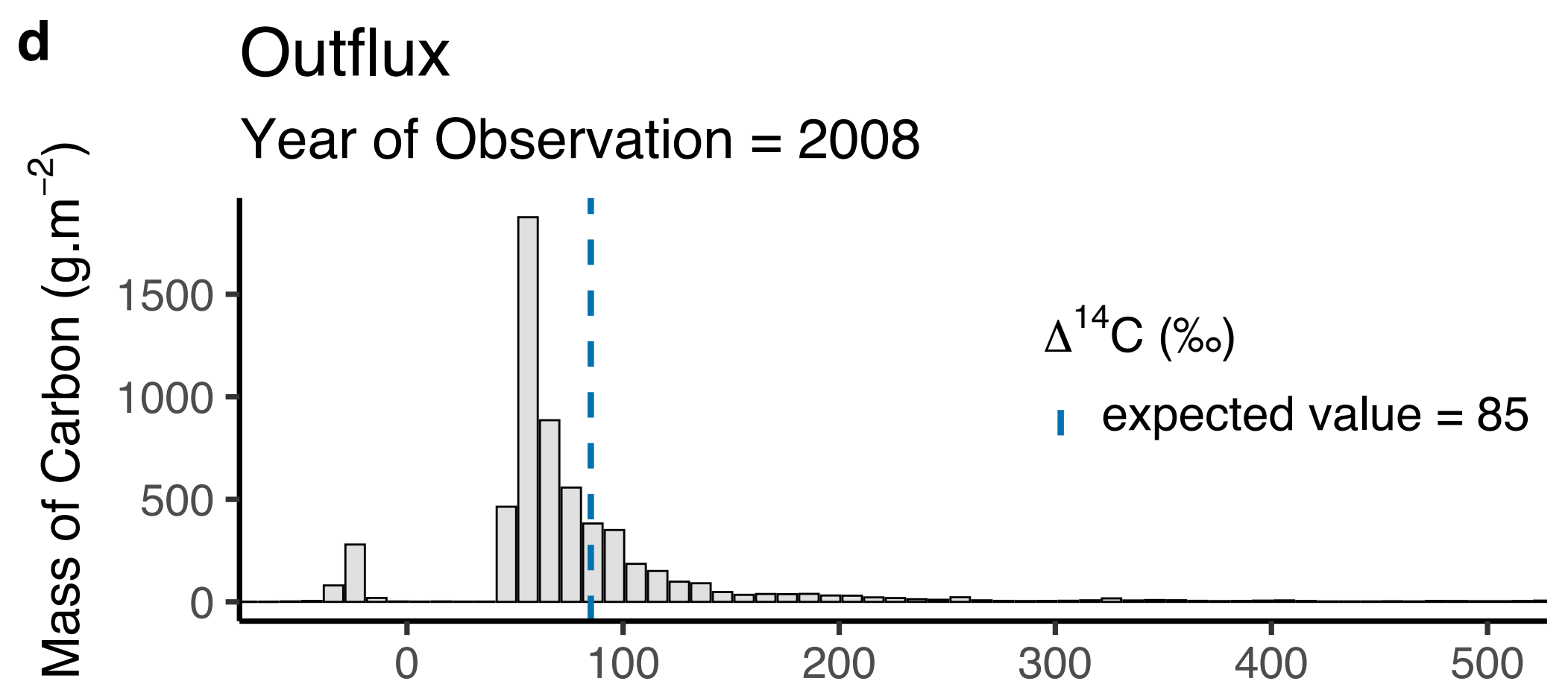
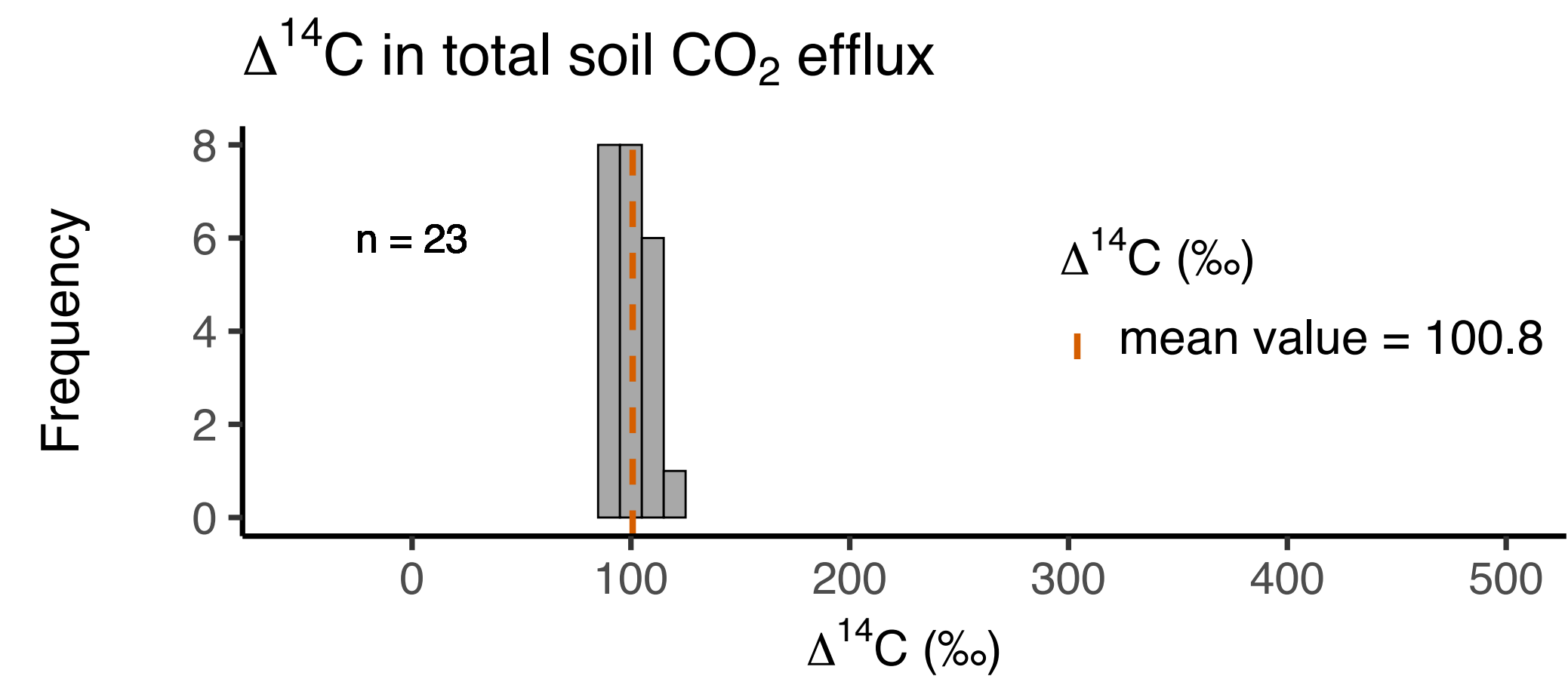
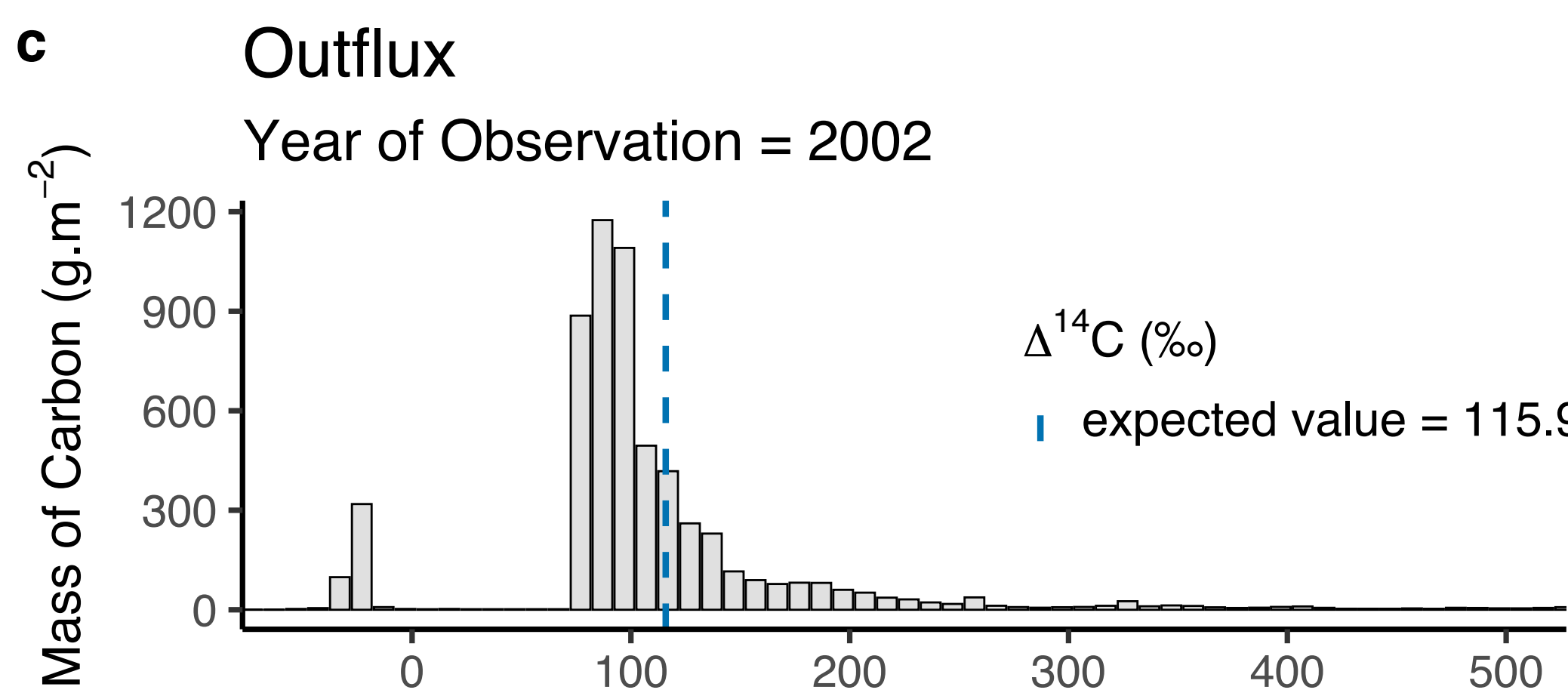
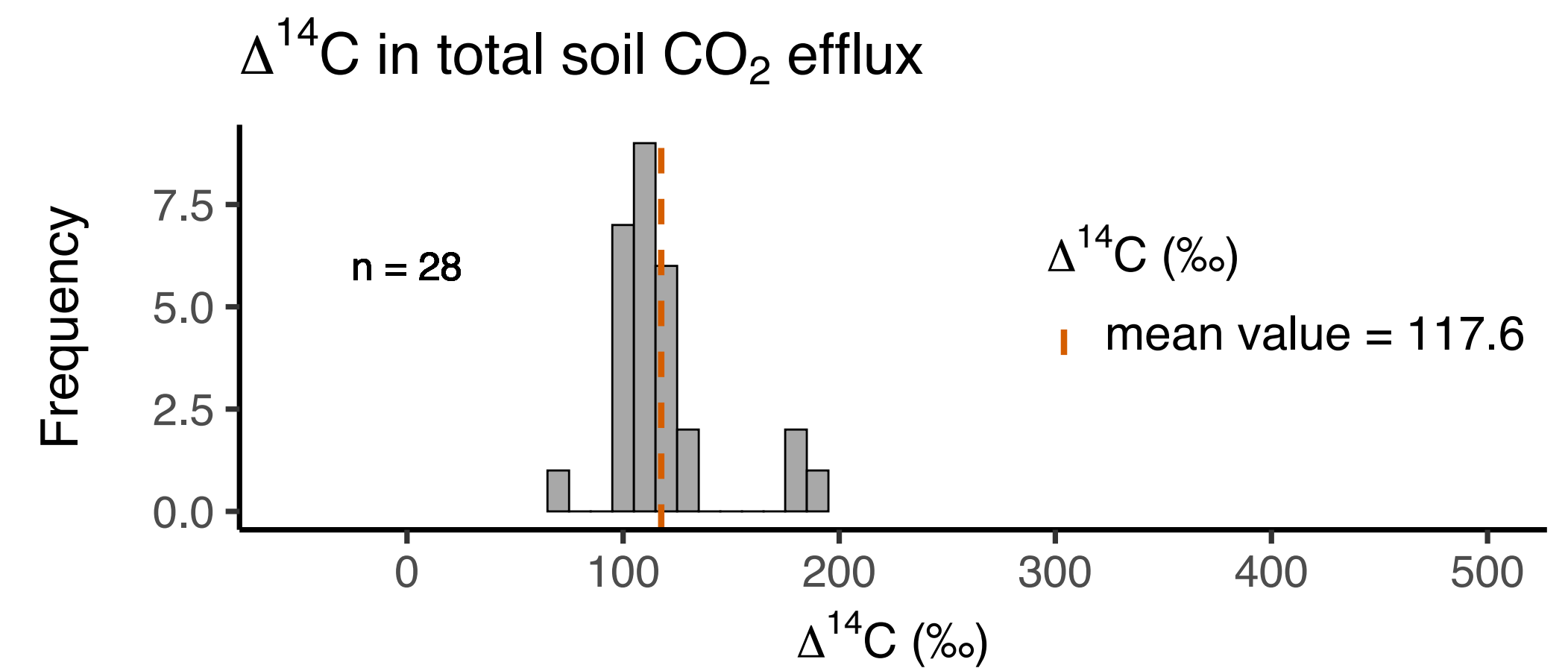
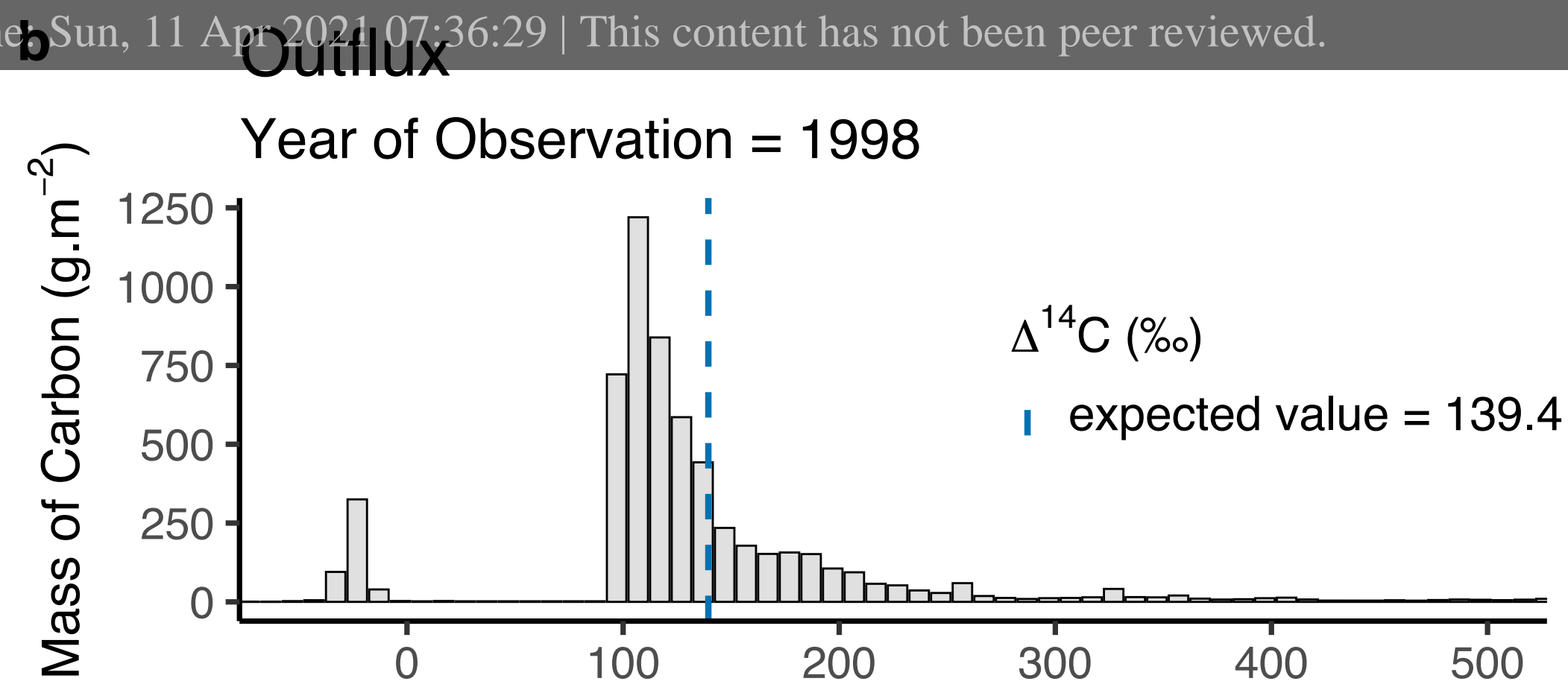
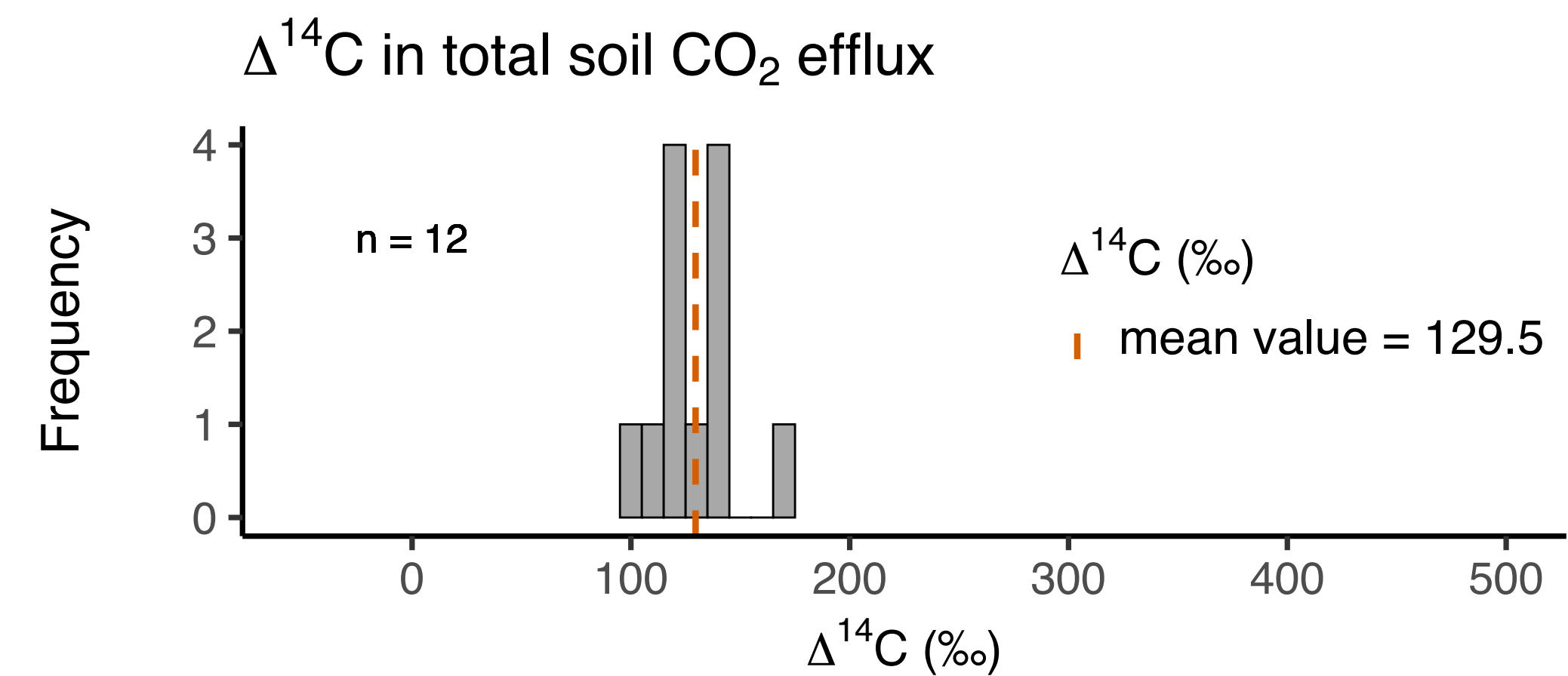
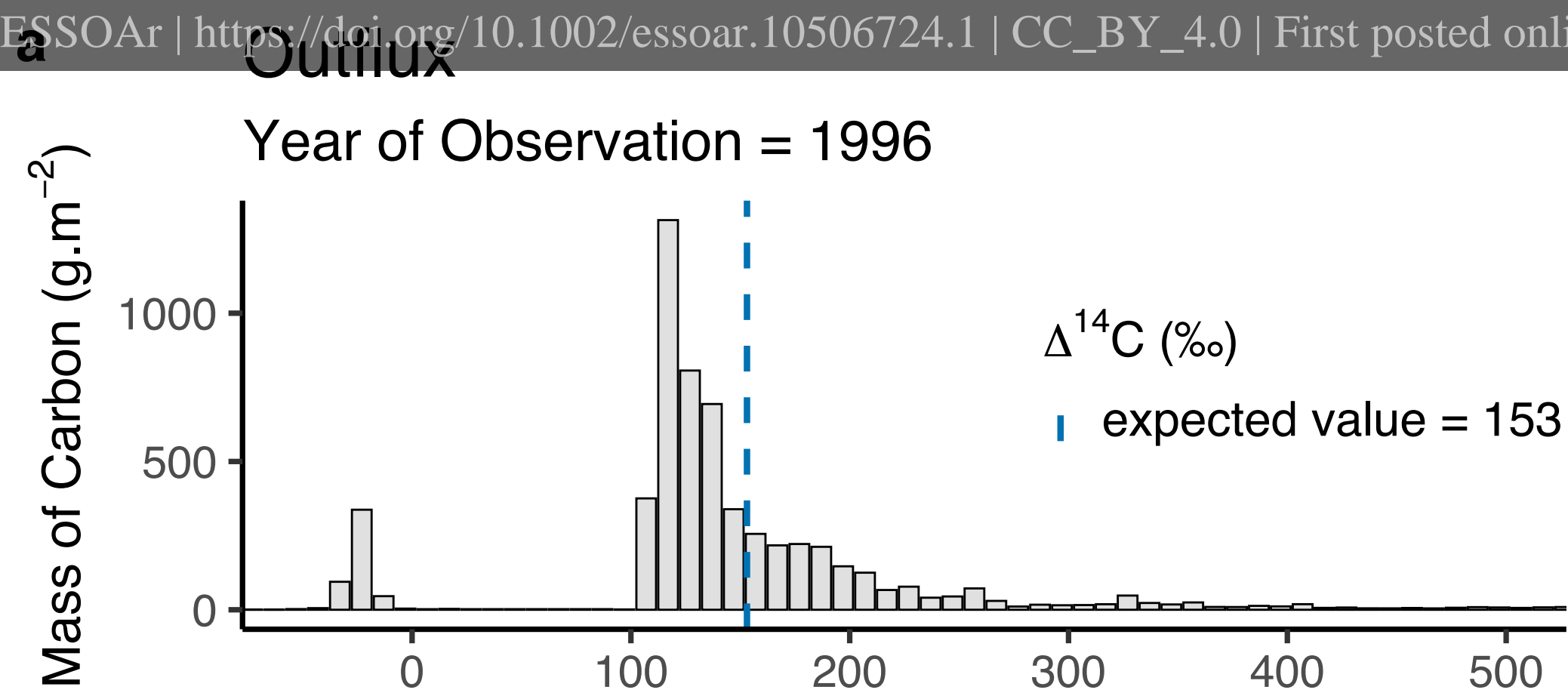
**Figure 8.**

# Evolution of expected $\Delta^{14}\text{C}$ over time



**Figure 9.**





**Figure 10.**

# Mean and Expected $\Delta^{14}\text{C}$ over time

Total Soil  $\text{CO}_2$  Efflux/Outflux

