Supporting Information

Metzler et al. 10.1073/pnas.1705296115

S1. The State Transition Matrix Φ

In contrast to the one-dimensional or the time-independent case, the state transition matrix for a time-dependent multidimensional system can in general not be computed analytically. It has nevertheless some useful properties, some of which we collect here. They can be found in refs. 24 and 25.

The state transition matrix of the system described by Eq. **3** is the solution of the matrix equation

$$\begin{aligned} \frac{d}{dt}\Phi(t,t_0) &= \mathrm{B}(t)\,\Phi(t,t_0), \quad t > t_0, \\ \Phi(t_0,t_0) &= \mathrm{I}, \end{aligned} \qquad \textbf{[S1.1]}$$

and is given by the Peano–Baker series

$$\begin{aligned} \Phi(t,t_0) = \mathrm{I} &+ \int_{t_0}^{t} \mathrm{B}(\tau_1)\,d\tau_1 \\ &+ \int_{t_0}^{t} \mathrm{B}(\tau_1) \int_{t_0}^{\tau_1} \mathrm{B}(\tau_2)\,d\tau_2\,d\tau_1 + \cdots . \end{aligned} \qquad \textbf{[S1.2]}$$

Here, $\mathrm{I} \in \mathbb{R}^{d\times d}$ is the identity matrix.

If $\mathrm{B}(t) = b(t)$ is a scalar, then the Peano–Baker series can be summed to

$$\Phi(t,t_0) = \exp\left(\int_{t_0}^{t} b(\tau)\,d\tau\right). \qquad \textbf{[S1.3]}$$

If $\mathrm{B}(t) = \mathrm{B}$ is a real constant $d \times d$ matrix, then

$$\begin{aligned} \Phi(t,t_0) &= \mathrm{I} + \frac{1}{1!}(t-t_0)^1\,\mathrm{B}^1 + \frac{1}{2!}(t-t_0)^2\,\mathrm{B}^2 + \cdots \\ &= e^{(t-t_0)\,\mathrm{B}}, \end{aligned} \qquad \textbf{[S1.4]}$$

where for a $d \times d$ matrix Q the expression e^{Q} denotes the matrix exponential. In many cases this matrix exponential can be computed explicitly. If further B is compartmental and invertible, then $(e^{(t-t_0)\,\mathrm{B}})_{t\geq t_0}$ is a semigroup of contractions, meaning that

$$\left\| e^{(t-t_0)\,\mathrm{B}}\,\mathbf{u} \right\| \leq e^{-\lambda\,(t-t_0)}\ \|\mathbf{u}\|, \quad t \geq t_0, \mathbf{u} \in \mathbb{R}^d. \qquad \textbf{[S1.5]}$$

Here, $\lambda > 0$ is such that $-\lambda$ is the largest real part of the eigenvalues of B, and the norm $\|\,\mathbf{v}\,\|$ of a vector $\mathbf{v} \in \mathbb{R}^d$ is given by

$$\|\mathbf{v}\| = \sum_{i=1}^{d} |v_i|. \qquad \textbf{[S1.6]}$$

More information on matrix exponentials and semigroups can be found in ref. 26.

If $\mathrm{B}(t) = b(t)\,\mathrm{M}$ is a scalar multiplied with a constant matrix, then

$$\Phi(t,t_0) = \exp\left(\int_{t_0}^{t} b(\tau)\,d\tau\,\mathrm{M}\right). \qquad \textbf{[S1.7]}$$

If $\mathrm{B}(t)$ is compartmental for all $t \geq t_0$, then its logarithmic norm $\mu(\mathrm{B}(t))$ is nonpositive. Consequently,

$$\sup_{\|\mathbf{u}\|=1} \|\Phi(t,t_0)\,\mathbf{u}\| \leq \exp\left(\int_{t_0}^{t} \mu(\mathrm{B}(\tau))\,d\tau\right) \leq 1. \qquad \textbf{[S1.8]}$$

S2. The McKendrick–von Foerster Equation

Eq. **8** can be interpreted as a multidimensional McKendrick–von Foerster equation, because for the ith compartment

$$\left(\frac{\partial}{\partial a} + \frac{\partial}{\partial t}\right) p_i(a,t) = \gamma_i(t)\,p_i(a,t), \qquad \textbf{[S2.1]}$$

where $\gamma_i(t) = \sum_{j\neq i} B_{ij}(t) + B_{ii}(t)$ is the combination of the incoming and outgoing rates of mass with age a at time t.

We prove now that our density function satisfies the McKendrick–von Foerster Eq. **8**. To that end, we compute the total differential of the density function along the characteristics $a(t) = a^0 + t$ by

$$\begin{aligned} \frac{d}{dt}\,\mathbf{p}(a,t) &= \frac{\partial}{\partial a}\,\mathbf{p}(a,t)\,\frac{d}{dt}\,a(t) + \frac{\partial}{\partial t}\,\mathbf{p}(a,t)\,\frac{d}{dt}\,t \\ &= \left(\frac{\partial}{\partial a} + \frac{\partial}{\partial t}\right)\,\mathbf{p}(a,t), \end{aligned} \qquad \textbf{[S2.2]}$$

where $a^0 \geq 0$ is some initial age. We continue in two steps. In the first step, we show that Eq. **8** holds on $S_1 := \{(a,t) : t \geq t_0, a \geq t - t_0\}$ with initial condition **10**. In the second step, we show that Eq. **8** holds on $S_2 := \{(a,t) : a \geq 0, t \geq t_0, a < t - t_0\}$ with boundary condition **9**.

A. Step 1. On S_1 we have $\mathbf{p}(a,t) = \mathbf{g}(a,t)$. Consequently, we prove the initial condition **10** by

$$\mathbf{p}(a,t_0) = \Phi(t_0,t_0)\,\mathbf{p^0}(a-(t_0-t_0)) = \mathbf{p^0}(a). \qquad \textbf{[S2.3]}$$

Furthermore,

$$\mathbf{p}(a,t) = \Phi(t,t_0)\,\mathbf{p^0}\left(a^0\right), \qquad \textbf{[S2.4]}$$

where $a^0 = a - (t - t_0)$ does not change with time on the characteristics. Consequently,

$$\begin{aligned} \left(\frac{\partial}{\partial a} + \frac{\partial}{\partial t}\right)\,\mathbf{p}(a,t) &= \frac{d}{dt}\,\mathbf{p}(a,t) \\ &= \frac{d}{dt}\,\Phi(t,t_0)\,\mathbf{p^0}\left(a^0\right) \\ &= \mathrm{B}(t)\,\Phi(t,t_0)\,\mathbf{p^0}\left(a^0\right) \\ &= \mathrm{B}(t)\,\mathbf{p}(a,t), \end{aligned} \qquad \textbf{[S2.5]}$$

which proves Eq. **8** on S_1.

B. Step 2. On S_2 we have $\mathbf{p}(a,t) = \mathbf{h}(a,t)$ and $a^0 = 0$. Consequently, we prove the boundary condition **9** by

$$\mathbf{p}(0,t) = \Phi(t,t-0)\,\mathbf{u}(t-0) = \mathbf{u}(t). \qquad \textbf{[S2.6]}$$

Furthermore,

$$\mathbf{p}(a,t) = \Phi(t,s)\,\mathbf{u}(s), \qquad \textbf{[S2.7]}$$

where $s = t - a$ does not change with time on the characteristics. Consequently,

$$\begin{aligned} \left(\frac{\partial}{\partial a} + \frac{\partial}{\partial t}\right)\,\mathbf{p}(a,t) &= \frac{d}{dt}\,\mathbf{p}(a,t) \\ &= \frac{d}{dt}\,\Phi(t,s)\,\mathbf{u}(s) \\ &= \mathrm{B}(t)\,\Phi(t,s)\,\mathbf{u}(s) \\ &= \mathrm{B}(t)\,\mathbf{p}(a,t), \end{aligned} \qquad \textbf{[S2.8]}$$

which proves Eq. **8** on S_2.

S3. The Semiexplicit Formula for Compartment Age Moments

We assume that the initial age density $\mathbf{p^0}$ admits finite moments up to order n and denote them by $\bar{\mathbf{a}}^{0,k}$, $k = 1, 2, \ldots, n$. Additionally, we define

$$\mathbf{y}(t) := \Phi(t, t_0)\, \mathbf{x^0}, \quad t \geq t_0, \qquad \textbf{[S3.1]}$$

and

$$\mathbf{z}(t) := \int_{t_0}^{t} \Phi(t, \tau)\, \mathbf{u}(\tau)\, d\tau, \quad t \geq t_0. \qquad \textbf{[S3.2]}$$

Consequently, $\mathbf{x} = \mathbf{y} + \mathbf{z}$, where $\mathbf{y}$ describes the evolution of the initial mass and $\mathbf{z}$ describes the evolution of mass that comes later into the system. We use the shorthand $\bar{\mathbf{a}}^n$ for $\bar{\mathbf{a}}^n(t) := \bar{\mathbf{a}}^{\mathbf{x}(t),n}$ and note that we can compute the nth moment of the age density of $\mathbf{x}$ by the corresponding moments of the age densities of $\mathbf{y}$ and $\mathbf{z}$ by

$$\bar{a}_i^n = \frac{y_i\, \bar{a}_i^{\mathbf{y},n} + z_i\, \bar{a}_i^{\mathbf{z},n}}{x_i}, \quad i = 1, 2, \ldots, d, \qquad \textbf{[S3.3]}$$

or, in vector notation,

$$\bar{\mathbf{a}}^n(t) = \mathrm{X}(t)^{-1} \left[\mathrm{Y}(t)\, \bar{\mathbf{a}}^{\mathbf{y},n}(t) + \mathrm{Z}(t)\, \bar{\mathbf{a}}^{\mathbf{z},n}(t) \right]. \qquad \textbf{[S3.4]}$$

We see from Eq. **15** that

$$\mathrm{Y}(t)\, \bar{\mathbf{a}}^{\mathbf{y},n}(t) = \int_0^\infty a^n\, \mathbf{g}(a, t)\, da, \qquad \textbf{[S3.5]}$$

which can be transformed by

$$\mathbf{g}(a, t) = \mathbb{1}_{[t-t_0,\infty)}(a)\, \Phi(t, t_0)\, \mathbf{p^0}(a - (t - t_0)) \qquad \textbf{[S3.6]}$$

and a change of variables from a to $\tau := a - (t - t_0)$ into

$$\mathrm{Y}(t)\, \bar{\mathbf{a}}^{\mathbf{y},n}(t) = \Phi(t, t_0) \int_0^\infty [\tau + (t - t_0)]^n\, \mathbf{p^0}(\tau)\, d\tau. \qquad \textbf{[S3.7]}$$

An application of the binomial theorem and Eq. **15** leads to

$$\mathrm{Y}(t)\, \bar{\mathbf{a}}^{\mathbf{y},n}(t) = \sum_{k=0}^{n} \binom{n}{k} (t - t_0)^{n-k}\, \Phi(t, t_0)\, \mathrm{X}^0\, \bar{\mathbf{a}}^{0,k}. \qquad \textbf{[S3.8]}$$

Furthermore, again by Eq. **15**,

$$\mathrm{Z}(t)\, \bar{\mathbf{a}}^{\mathbf{z},n}(t) = \int_0^\infty a^n\, \mathbf{h}(a, t)\, da = \int_0^{t-t_0} a^n\, \mathbf{h}(a, t)\, da. \qquad \textbf{[S3.9]}$$

We plug the sum of Eqs. **S3.8** and **S3.9** into Eq. **S3.4** to establish Eq. **16**.

S4. The Compartment Age Moment System

We assume that the initial age density $\mathbf{p^0}$ admits finite moments up to order n and denote them by $\bar{\mathbf{a}}^{0,k}$, $k = 1, 2, \ldots, n$. Furthermore, we assume $\mathbf{y}$ and $\mathbf{z}$ as in Eqs. **S3.1** and **S3.2**, respectively.

Our goal is to derive a system of ODEs for the moments up to order n for the age distributions of $\mathbf{x}$. To this end, we try to represent the time derivative of the kth moment by known quantities. For that purpose, we need some auxiliary results.

A. Auxiliary Results.

Lemma S4.1. *For $k = 1, 2, \ldots, n$, and $t > t_0$,*

$$\frac{d}{dt} \int_0^\infty a^k\, g_i(a, t)\, da = \sum_{j=1}^{d} B_{ij}(t)\, y_j(t)\, \bar{a}_j^{\mathbf{y},k}(t) + k\, y_i(t)\, \bar{a}_i^{\mathbf{y},k-1}(t). \qquad \textbf{[S4.1]}$$

Proof: For simplicity of notation, we do not consider the component i, but the entire vector. We start with

$$\frac{d}{dt} \int_0^\infty a^k\, \mathbf{g}(a, t)\, da \qquad \textbf{[S4.2]}$$

and use

$$\mathbf{g}(a, t) = \mathbb{1}_{[t-t_0,\infty)}\, \Phi(t, t_0)\, \mathbf{p^0}(a - (t - t_0)) \qquad \textbf{[S4.3]}$$

to obtain

$$\frac{d}{dt} \Phi(t, t_0) \int_{t-t_0}^\infty a^k\, \mathbf{p^0}(a - (t - t_0))\, da, \qquad \textbf{[S4.4]}$$

which by the product rule turns into

$$\mathrm{B}(t)\, \Phi(t, t_0) \int_{t-t_0}^\infty a^k\, \mathbf{p^0}(a - (t - t_0))\, da + \Phi(t, t_0) \frac{d}{dt} \int_{t-t_0}^\infty a^k\, \mathbf{p^0}(a - (t - t_0))\, da. \qquad \textbf{[S4.5]}$$

We transform the first term back. In addition, a change of variables in the second term from a to $\tau := a - (t - t_0)$ brings

$$\mathrm{B}(t) \int_0^\infty a^k\, \mathbf{g}(a, t)\, da + \Phi(t, t_0) \frac{d}{dt} \int_0^\infty (\tau + (t - t_0))^k\, \mathbf{p^0}(\tau)\, d\tau. \qquad \textbf{[S4.6]}$$

We use Eq. **15** in the first term and in the second term we bring the derivative under the integral by means of the dominated convergence theorem to get

$$\mathrm{B}(t)\, \mathrm{Y}(t)\, \bar{\mathbf{a}}^{\mathbf{y},k}(t) + \Phi(t, t_0) \int_0^\infty k\, (\tau + (t - t_0))^{k-1}\, \mathbf{p^0}(\tau)\, d\tau. \qquad \textbf{[S4.7]}$$

We undo the change of variables in the second term and transform it back to obtain

$$\mathrm{B}(t)\, \mathrm{Y}(t)\, \bar{\mathbf{a}}^{\mathbf{y},k}(t) + k \int_{t-t_0}^\infty a^{k-1}\, \mathbf{g}(a, t)\, da, \qquad \textbf{[S4.8]}$$

which equals

$$\mathrm{B}(t)\, \mathrm{Y}(t)\, \bar{\mathbf{a}}^{\mathbf{y},k}(t) + k\, \mathrm{Y}(t)\, \bar{\mathbf{a}}^{\mathbf{y},k-1}. \qquad \textbf{[S4.9]}$$

Computing the ith component, we get

$$\sum_{j=1}^{d} B_{ij}(t)\, y_j(t)\, \bar{a}_j^{\mathbf{y},k} + k\, y_i(t)\, \bar{a}_i^{\mathbf{y},k-1}(t) \qquad \textbf{[S4.10]}$$

and we are finished with the proof.

□

Lemma S4.2. *For $k = 1, 2, \ldots, n$, and $t > t_0$,*

$$\frac{d}{dt} \int_0^\infty a^k\, h_i(a, t)\, da = \sum_{j=1}^{d} B_{ij}(t)\, z_j(t)\, \bar{a}_i^{\mathbf{z},k}(t) + k\, z_i(t)\, \bar{a}_i^{\mathbf{z},k-1}(t). \qquad \textbf{[S4.11]}$$

Proof: Again, for simplicity of notation, we do not consider the component i, but the entire vector. From

$$\mathbf{h}(a, t) = \mathbb{1}_{[0,t-t_0)}(a)\, \Phi(t, t - a)\, \mathbf{u}(t - a), \qquad \textbf{[S4.12]}$$

we get

$$\int_0^\infty a^k\,\mathbf{h}(a,t)\,da = \lim_{\varepsilon\searrow 0}\int_0^{t-t_0-\varepsilon} a^k\,\mathbf{h}(a,t)\,da. \quad \textbf{[S4.13]}$$

We can interchange the limit and the derivative to see

$$\frac{d}{dt}\int_0^\infty a^k\,\mathbf{h}(a,t)\,da = \lim_{\varepsilon\searrow 0}\frac{d}{dt}\int_0^{t-t_0-\varepsilon} a^k\,\mathbf{h}(a,t)\,da. \quad \textbf{[S4.14]}$$

By an application of the Leibniz rule to the right-hand side we obtain

$$\lim_{\varepsilon\searrow 0}\int_0^{t-t_0-\varepsilon} a^k\,\frac{\partial}{\partial t}\mathbf{h}(a,t)\,da + (t-t_0-\varepsilon)^k\,\mathbf{h}(t-t_0-\varepsilon,t). \quad \textbf{[S4.15]}$$

In *S2. The McKendrick–von Foerster Equation, B. Step 2* we derived that for $a\in[0,t-t_0-\varepsilon]$,

$$\frac{\partial}{\partial t}\mathbf{h}(a,t) = \mathrm{B}(t)\,\mathbf{h}(a,t) - \frac{\partial}{\partial a}\mathbf{h}(a,t), \quad \textbf{[S4.16]}$$

which we plug into the first term of Eq. **S4.15** and turn it into

$$\lim_{\varepsilon\searrow 0}\int_0^{t-t_0-\varepsilon} a^k\left[\mathrm{B}(t)\,\mathbf{h}(a,t) - \frac{\partial}{\partial a}\mathbf{h}(a,t)\right]da, \quad \textbf{[S4.17]}$$

which equals by Eq. **15**

$$\mathrm{B}(t)\,\mathrm{Z}(t)\,\bar{\mathbf{a}}^{\mathbf{z},k} - \lim_{\varepsilon\searrow 0}\int_0^{t-t_0-\varepsilon} a^k\,\frac{\partial}{\partial a}\mathbf{h}(a,t)\,da. \quad \textbf{[S4.18]}$$

We integrate by parts and use again Eq. **15** to get

$$\begin{aligned}\mathrm{B}(t)\,\mathrm{Z}(t)\,\bar{\mathbf{a}}^{\mathbf{z},k} &- \lim_{\varepsilon\searrow 0}(t-t_0-\varepsilon)^k\,\mathbf{h}(t-t_0-\varepsilon,t)\\ &+ k\,\mathrm{Z}(t)\,\bar{\mathbf{a}}^{\mathbf{z},k-1}.\end{aligned} \quad \textbf{[S4.19]}$$

Together with Eq. **S4.15**, we have

$$\frac{d}{dt}\int_0^\infty a^k\,\mathbf{h}(a,t)\,da = \mathrm{B}(t)\,\mathrm{Z}(t)\,\bar{\mathbf{a}}^{\mathbf{z},k} + k\,\mathrm{Z}(t)\,\bar{\mathbf{a}}^{\mathbf{z},k-1}, \quad \textbf{[S4.20]}$$

which completes the proof by considering the ith component. □

Lemma S4.3. *For* $k=1,2,\dots,n$, *and* $t>t_0$,

$$\begin{aligned}\frac{d}{dt}\left(x_i(t)\,\bar{a}_i^k(t)\right) &= \sum_{j=1}^d B_{ij}(t)\,x_j(t)\,\bar{a}_j^k(t)\\ &+ k\,x_i(t)\,\bar{a}_i^{k-1}(t).\end{aligned} \quad \textbf{[S4.21]}$$

Proof: From Eq. **15** and $\mathbf{p}(a,t)=\mathbf{g}(a,t)+\mathbf{h}(a,t)$, we know

$$\begin{aligned}\frac{d}{dt}\left[x_i(t)\,\bar{a}_i^k(t)\right] &= \frac{d}{dt}\int_0^\infty a^k\,p_i(a,t)\,da\\ &= \frac{d}{dt}\int_0^\infty a^k\,g_i(a,t)\,da + \frac{d}{dt}\int_0^\infty a^k\,h_i(a,t)\,da.\end{aligned} \quad \textbf{[S4.22]}$$

Consequently, we can apply *Lemmas S4.1* and *S4.2* and use

$$x_i\,\bar{a}_i{}^k = y_i\,\bar{a}_i^{\mathbf{y},k} + z_i\,\bar{a}_i^{\mathbf{z},k} \quad \textbf{[S4.23]}$$

from Eq. **S3.3** to obtain

$$\begin{aligned}\frac{d}{dt}\left(x_i\,\bar{a}_i^k\right) &= \sum_{j=1}^d B_{ij}\,y_j\,\bar{a}_j^{\mathbf{y},k} + k\,y_i\,\bar{a}_i^{\mathbf{y},k-1}\\ &\quad + \sum_{j=1}^d B_{ij}\,z_j\,\bar{a}_j^{\mathbf{z},k} + k\,z_i\,\bar{a}_i^{\mathbf{z},k-1}\\ &= \sum_{j=1}^d B_{ij}\left(y_j\,\bar{a}_j^{\mathbf{y},k} + z_j\,\bar{a}_j^{\mathbf{z},k}\right)\\ &\quad + k\left(y_j\,\bar{a}_i^{\mathbf{y},k-1} + z_j\,\bar{a}_i^{\mathbf{z},k-1}\right)\\ &= \sum_{j=1}^d B_{ij}\,x_j\,\bar{a}_j^k + k\,\bar{a}_i^{k-1}.\end{aligned} \quad \textbf{[S4.24]}$$

□

B. Proof of the Compartment Age Moment System. Let $k\in\{1,2,\dots,n\}$. We compute the time derivative of $\bar{a}_i^k$ at $t>t_0$ by

$$\frac{d}{dt}\bar{a}_i^k(t) = \frac{d}{dt}\left[\frac{x_i(t)\,\bar{a}_i^k(t)}{x_i(t)}\right]. \quad \textbf{[S4.25]}$$

We apply the quotient rule and *Lemma S4.3* to get

$$\begin{aligned}\frac{d}{dt}\bar{a}_i^k &= \frac{1}{x_i^2}\left[\left(\sum_{j=1}^d B_{ij}\,x_j\,\bar{a}_j^k + k\,x_i\,\bar{a}_i^{k-1}\right)x_i \right.\\ &\qquad\left. - x_i\,\bar{a}_i^k\,\frac{d}{dt}x_i\right]\\ &= k\,\bar{a}_i^{k-1} + \frac{1}{x_i}\left[\sum_{j=1}^d B_{ij}\,x_j\,\bar{a}_j^k \right.\\ &\qquad\left. - \bar{a}_i^k\left(\sum_{j=1}^d B_{ij}\,x_j + u_i\right)\right]\\ &= k\,\bar{a}_i^{k-1}\\ &\quad + \frac{1}{x_i}\left[\sum_{j=1}^d B_{ij}\,x_j\left(\bar{a}_j^k - \bar{a}_i^k\right) - \bar{a}_i^k\,u_i\right].\end{aligned} \quad \textbf{[S4.26]}$$

Now, we can bring all components $i=1,2,\dots,d$ into one vector and the proof is complete.

S5. The Age Quantiles

We want to show that the compartment age quantiles solve the initial value problem given by Eq. **21**. The time evolution of the kth n quantile $\xi_i(t)$ of the age of compartment i can be described by taking the time derivative in both sides of Eq. **20**, which gives

$$\frac{d}{dt}\int_0^{\xi_i(t)} p_i(a,t)\,da = q\,[\mathrm{B}(t)\,\mathbf{x}(t)]_i + q\,u_i(t). \quad \textbf{[S5.1]}$$

Using the Leibniz rule, we can rewrite the left-hand side to

$$\int_0^{\xi_i(t)} \frac{\partial}{\partial t}p_i(a,t)\,da + p_i(\xi_i(t),t)\,\frac{d}{dt}\xi_i(t). \quad \textbf{[S5.2]}$$

Outside the null set $\{a \geq 0: \; a = t - t_0\}$ the McKendrick–von Foerster Eq. **8** holds. Consequently,

$$\int_0^{\xi_i(t)} \frac{\partial}{\partial t}\, p_i(a,t)\, da = \int_0^{\xi_i(t)} \left([\mathrm{B}(t)\, \mathbf{p}(a,t)]_i - \frac{\partial}{\partial a}\, p_i(a,t) \right) da, \qquad \textbf{[S5.3]}$$

which turns by integration by parts and $p_i(0,t) = u_i(t)$ into

$$[\mathrm{B}(t)\, \mathbf{P}(\xi_i(t), t)]_i - p_i(\xi_i(t), t) + u_i(t). \qquad \textbf{[S5.4]}$$

We plug it into Eq. **S5.2** and the result into Eq. **S5.1** to prove Eq. **21**. The initial value ξ_i^0 is just the kth n quantile of the initial age of the content of compartment i. The proof for the system age quantile ξ follows analogously from $\| \mathbf{P}(\xi(t), t) \| = q \; \| \mathbf{x}(t) \|$.

S6. The Relations to Earlier Results and Possible Applications

A. Linear Time-Independent Systems. Our results generalize earlier results on explicit formulas for ages and transit times of linear time-independent compartmental systems in steady state. To arrive at this conclusion, we consider such a system given by

$$\frac{d}{dt}\mathbf{x}(t) = \mathrm{B}\,\mathbf{x}(t) + \mathbf{u}, \quad t > t_0, \qquad \mathbf{x}(t_0) = \mathbf{x^0}, \qquad \textbf{[S6.1]}$$

where B is an invertible compartmental matrix. Recall that in time-independent systems both B and $\mathbf{u}$ are independent of time. The state transition matrix $\Phi(t,s) = e^{(t-s)\,\mathrm{B}}$ is then given by the matrix exponential. Therefore, for $a \geq 0$ and $t \geq t_0$,

$$\mathbf{p}(a,t) = e^{(t-t_0)\,\mathrm{B}}\, \mathbf{p^0}(a - (t - t_0)) + e^{a\,\mathrm{B}}\, \mathbf{u}. \qquad \textbf{[S6.2]}$$

Since $\lim_{t\to\infty} e^{t\,\mathrm{B}} = 0$ if B is compartmental and invertible, $\lim_{t\to\infty} \mathbf{p}(a,t) = e^{a\,\mathrm{B}}\,\mathbf{u}$. This vector contains the compartment age densities of system **S6.1** after it has run for an infinite time. Hence, they belong to the steady state $\mathbf{x}^* = -\mathrm{B}^{-1}\,\mathbf{u}$ of the system. If we divide each component of the steady-state age density vector by the steady-state content of the corresponding compartment, we obtain the normalized age density vector from ref. 12; namely

$$\mathbf{f_a}(a) = (\mathrm{X}^*)^{-1}\, e^{a\,\mathrm{B}}\,\mathbf{u}, \quad a \geq 0, \qquad \textbf{[S6.3]}$$

where $\mathrm{X}^* = \mathrm{diag}(x_1^*, x_2^*, \ldots, x_d^*)$. This means that along with the compartment contents also the age distributions converge to their steady state as $t \to \infty$.

If we start the system with the steady-state age structure by choosing $\mathbf{p^0}(a) = e^{a\,\mathrm{B}}\,\mathbf{u}$, then

$$\mathbf{x^0} = \int_0^{\infty} e^{a\,\mathrm{B}}\,\mathbf{u}\, da = -\mathrm{B}^{-1}\,\mathbf{u} = \mathbf{x}^*. \qquad \textbf{[S6.4]}$$

For $t \geq t_0$ and $a > t - t_0$ we have

$$\mathbf{p}(a,t) = e^{(t-t_0)\,\mathrm{B}}\, e^{[a-(t-t_0)]\,\mathrm{B}}\,\mathbf{u} = e^{a\,\mathrm{B}}\,\mathbf{u}, \qquad \textbf{[S6.5]}$$

and for $t \geq t_0$ and $a \leq t - t_0$ we have also

$$\mathbf{p}(a,t) = e^{[t-(t-a)]\,\mathrm{B}}\,\mathbf{u} = e^{a\,\mathrm{B}}\,\mathbf{u}. \qquad \textbf{[S6.6]}$$

Consequently, both the system's content and its age structure remain constant for all time if the system is in steady state.

Since the backward transit time is the age of a particle at the moment when it leaves the system, in steady state and after normalization, Eq. **24** coincides with the formula given in ref. 12. The same formula holds also for the forward transit-time density in steady state, since by Eq. **28** the forward transit time is only a time-shifted backward transit time.

B. Different Approaches for Different Scenarios. Depending on the structure of the system, there exist many different approaches to obtain transit-time and age distributions, most of which are special cases of our present results.

B.1. Linear time-independent systems. If the compartmental system is linear and time independent, then the response function approach is very useful to establish formulas for transit-time and age densities. A first step in this direction was done by computations depending on a system response function, which was not explicitly known (9). This system response function returns the proportion $h(\tau)$ of the input that leaves the system when time τ has passed by. The concept of response functions was also the basis to obtain the desired densities numerically by long-term simulations in two carbon-cycle models by computing the system response to impulsive inputs (11). The resulting impulse response function $\boldsymbol{\psi}$ depends in the first place on the fixed and constant impulse. Then $\boldsymbol{\psi}(\tau)$ is the vector of mass leaving the system after time τ has elapsed, where each component of the vector belongs to a compartment. The impulse response approach was later investigated theoretically to obtain transit-time and age densities explicitly for a set of carbon-cycle models of very simple structure, using Laplace transforms (10). Eventually, both the normalized transit-time and the normalized age density are simply probability density functions of a phase-type distribution, and the impulse response function equals the normalized transit-time density (12). If furthermore the corresponding compartmental matrix $\mathrm{B} \in \mathbb{R}^{d\times d}$ is invertible, which holds true for trap-free open systems (6), then by Eq. **S1.5** it is obvious that the age densities decay exponentially. The exponential decay rate λ corresponds to the e-folding time of the longest-lived mode of stratospheric transport (27).

Consequently, our present work generalizes the response function approach by allowing for time-dependent parameters and nonlinear dependencies.

B.2. Linear time-dependent systems. Response theory for linear systems with time-dependent parameters has been present for a long time (28). Our present work generalizes those results to a multidimensional, possibly nonlinear setting. In addition, we provide semiexplicit formulas to compute the time-dependent system response function.

B.3. Green's Function. The Green's function approach is very common, for example, in atmospheric sciences. Regarding the age of stratospheric air, the term "age spectrum" was coined for the age density of a fixed box in the stratosphere (29). Then the age spectrum was identified as a Green's function which governs the transport of particles from the tropical tropopause to the stratosphere (30). The stratosphere is modeled as $\mathbb{R}^n$, for $n = 1, 2, 3$. Mostly, the transport is considered to be stationary, which makes the Green's function depend only on one time variable. Consequently, the corresponding Green's function G belongs to a partial differential equation and $G(\tau, x_1, x_2)$ is the mass or concentration that moved from x_1 to x_2 in time τ. The main differences between that approach and ours are the different interpretations of space, since we consider the $\mathbb{R}^n$ discretized into d compartments. After this discretization, the Green's function for linear time-independent systems is a matrix exponential.

The age spectrum can also be considered as the transit-time probability density function from a region Ω to a point in space (31). This is in our context the age density and we point out that we instead denote as transit time the time a particle needs from its entry into the system until its exit. In ref. 31, the simplification to stationary transport is not required, and consequently the Green's function depends on two time variables, just like ours does. However, since we discretize the space into compartments, our Green's function Φ does not belong to a partial differential equation anymore, but to a linear ODE on $\mathbb{R}^d$. As a result, our Green's function is not scalar valued, but matrix valued. Such a Green's function is better known as a state transition matrix.

B.4. Nonlinear systems. A classical approach to treat nonlinear systems is the linearization of the system in the neighborhood of a fixed point (32). On the one hand, this approach has the advantage that now the simpler linear steady-state theory can be applied. On the other hand, the resulting outcomes are valid only in the vicinity of this particular fixed point and thus possibly only after an infinite amount of time has passed. Our approach, however, requires neither the existence of fixed points nor an infinite history.

B.5. Mixed compartments. A different approach from ours is needed when the well-mixed assumption of the compartments is dropped. The fluxes could be age dependent, which is a very common case in hydrology, where the focus mostly lies on the annual water balance of catchments (33). Such catchments are usually modeled as one compartment with one influx (precipitation) and two age-dependent outfluxes (evaporation, runoff) (13, 15, 34). Even though this case does not fit directly in our framework, it is possible to approximate the one-compartment system with age-dependent outflow by a multiple-compartment well-mixed system. For time-independent systems, this approximation is based on the fact that every nonnegative probability distribution can be approximated arbitrarily well by a phase-type distribution (35). Doing a similar kind of approximation for a time-dependent single-catchment model allows the full application of the theory presented here. A recent commentary emphasizes the restrictions of single-catchment models and highlights the need for splitting the single catchment into several compartments (ref. 33, p 396). Our results deliver the demanded "theoretical framework that includes both flow and the age distribution of these flowing and stored waters."

S7. The Detailed Model Description

The model consists of three compartments: atmosphere (A), terrestrial biosphere (T), and surface ocean (S). The letter D stands for the external compartment deep ocean with infinite content. We denote by $C_A = C_A(t)$, $C_T = C_T(t)$, and $C_S = C_S(t)$ the respective carbon contents in petagrams of carbon at time t in years. Two external fluxes add carbon to the system. The first one, u_S, is constant and goes from the deep ocean to the surface ocean, whereas the second one, $u_A = u_A(t)$, is time dependent and represents carbon added to the atmosphere by the burning of fossil fuels. Carbon can leave the system only if it moves from the surface ocean to the deep ocean. A flux from compartment X to compartment Y is denoted by F_{XY} and the following fluxes exist in the model, all given in $\mathrm{Pg\,C\,y^{-1}}$petagrams of carbon per year:

$$\begin{aligned} F_{AT} &= 60\,(C_A/700)^{\alpha}, & F_{AS} &= 100\,C_A/700, \\ F_{TA} &= 60\,C_T/3{,}000 + f_{TA}, & F_{SA} &= 100\,(C_S/1{,}000)^{\beta}, \\ F_{SD} &= 45\,C_S/1{,}000, & u_S &= 45. \end{aligned} \quad \textbf{[S7.1]}$$

Here, $f_{TA} = f_{TA}(t)$ represents an internal flux from the terrestrial biosphere to the atmosphere caused by land use change (deforestation). Its values and also the values of the external inputs through fossil fuel emissions, $u_A(t)$, are taken as time series data from the RCP/ECP8.5 scenario. The two parameters α and β control the fluxes from the atmosphere to the terrestrial biosphere and from the surface ocean to the atmosphere, respectively. If both parameters are equal to 1 and f_{TA} vanishes, then the model is linear; otherwise it is nonlinear.

The model can now be described by the three ODEs, for $t > t_0 = 1765$,

$$\begin{aligned} \frac{d}{dt}\,C_A(t) &= F_{TA}(t) + F_{SA}(t) - F_{AT}(t) - F_{AS}(t) + u_A(t), \\ \frac{d}{dt}\,C_T(t) &= F_{AT}(t) - F_{TA}(t), \\ \frac{d}{dt}\,C_S(t) &= F_{AS}(t) - F_{SA}(t) - F_{SD}(t) + u_S(t). \end{aligned} \quad \textbf{[S7.2]}$$

Note that the right-hand side of Eq. **S7.2** depends through Eq. **S7.1** not only on t, but also on the state vector $\mathbf{x}(t) = (C_A(t), C_T(t), C_S(t))^T$. If we now define the state- and time-dependent compartmental matrix $\mathrm{B} = \mathrm{B}(\mathbf{x}(t), t)$ to equal

$$\begin{pmatrix} -C_A^{-1}\,(F_{AT} + F_{AS}) & C_T^{-1}\,F_{TA} & C_S^{-1}\,F_{SA} \\ C_A^{-1}\,F_{AT} & -C_T^{-1}\,F_{TA} & 0 \\ C_A^{-1}\,F_{AS} & 0 & -C_S^{-1}\,(F_{SA} + F_{SD}) \end{pmatrix} \quad \textbf{[S7.3]}$$

and $\mathbf{u}(t) := (u_A(t), 0, u_S)^T$, then the model fits in the framework of Eq. **1** describing the nonlinear time-dependent compartmental system

$$\begin{aligned} \frac{d}{dt}\mathbf{x}(t) &= \mathrm{B}(\mathbf{x}(t), t)\,\mathbf{x}(t) + \mathbf{u}(t), \quad t > t_0, \\ \mathbf{x}(t_0) &= \mathbf{x^0}. \end{aligned} \quad \textbf{[S7.4]}$$

Since at time $t_0 = 1765$ the system is supposed to be in equilibrium,

$$\mathbf{x^0} = (700, 3{,}000, 1{,}000)^T. \quad \textbf{[S7.5]}$$

S8. The Derivation of the Results from the Example Application

First, we solve Eq. **S7.4** numerically on the time interval $[1765, 2500]$ and obtain a solution trajectory $\mathbf{x} = \mathbf{x}(t)$. With this solution in hand, we can at all times $t \in [1765, 2500]$ compute the compartmental matrix $\mathrm{B} = \mathrm{B}(\mathbf{x}(t), t)$.

A. Equilibrium Age Densities. At time $t_0 = 1765$ the system is supposed to be in equilibrium and the land use flux $f_{TA}(t_0)$ vanishes. We plug Eqs. **S7.1** and **S7.5** into Eq. **S7.3** and get

$$\mathrm{B}(\mathbf{x^0}, t_0) = \begin{pmatrix} -160/700 & 60/3{,}000 & 100/1{,}000 \\ 60/700 & -60/3{,}000 & 0 \\ 100/700 & 0 & -145/1{,}000 \end{pmatrix}. \quad \textbf{[S8.1]}$$

If we set $\mathrm{B}^0 := \mathrm{B}(\mathbf{x^0}, t_0)$ and $\mathbf{u^0} := \mathbf{u}(t_0) = (0, 0, 45)^T$, then $\mathrm{B}^0\,\mathbf{x^0} + \mathbf{u^0} = \mathbf{0}$. We further define $\mathrm{X}^0 = \mathrm{diag}(x_1^0, x_2^0, x_3^0)$ and apply the steady-state formula

$$\mathbf{p^0}(a) = \left(\mathrm{X}^0\right)^{-1} e^{a\,\mathrm{B}^0}\,\mathbf{u^0}, \quad a \geq 0, \quad \textbf{[S8.2]}$$

from ref. 12 to obtain the vector-valued function $\mathbf{p^0}$ of age densities in equilibrium.

B. Atmospheric Age. Fig. 4 depicts the 2D surface corresponding to $\mathbf{p} = \mathbf{p}(a, t)$ in the time interval 1765–2500 and the age interval 0–250 y. The scalar field $\mathbf{p}$ can be obtained by Eq. **5**. In Eq. **S8.2** we have already computed the initial age density $\mathbf{p^0}$, and the input vector $\mathbf{u}$ is given by the RCP/ECP8.5 scenario. Consequently, we are missing only the state transition matrix Φ. We compute Φ by numerically solving the matrix ODE **S1.1** on $\{(t_2, t_1) \in [1765, 2500] \times [1765, 2500] : t_2 \geq t_1\}$ and can then proceed to compute $\mathbf{p}$ on $[0, 250] \times [1765, 2500]$.

To obtain a time trajectory of the mean age and the second moment of the atmospheric carbon, we follow Eq. **17** and solve the nine-dimensional ODE system, for $t_0 = 1765$,

$$\frac{d}{dt}\begin{pmatrix}\mathbf{x}\\ \bar{\mathbf{a}}^1\\ \bar{\mathbf{a}}^2\end{pmatrix}(t) = \begin{pmatrix}\mathrm{B}(\mathbf{x}(t),t)\,\mathbf{x}(t)+\mathbf{u}(t)\\ \boldsymbol{\gamma}^1(t,\mathbf{x}(t),\mathbf{1},\bar{\mathbf{a}}^1(t))\\ \boldsymbol{\gamma}^2(t,\mathbf{x}(t),\bar{\mathbf{a}}^1(t),\bar{\mathbf{a}}^2(t))\end{pmatrix},\quad t>t_0,$$
$$\left(\mathbf{x},\bar{\mathbf{a}}^1,\bar{\mathbf{a}}^2\right)(t_0) = \left(\mathbf{x^0},\bar{\mathbf{a}}^{0,1},\bar{\mathbf{a}}^{0,2}\right), \qquad \textbf{[S8.3]}$$

where, for $k=1,2$, $\boldsymbol{\gamma}^k = (\gamma_1^k,\gamma_2^k,\gamma_3^k)^T$ and for $i=1,2,3,$

$$\gamma_i^k(t,\mathbf{x},\bar{\mathbf{a}}^{k-1},\bar{\mathbf{a}}^k) = k\,\bar{a}_i^{k-1} + \frac{1}{x_i}\left[\sum_{j=1}^{d} B_{ij}\,x_j\,\left(\bar{a}_j^k-\bar{a}_i^k\right)-\bar{a}_i^k\,u_i\right]. \qquad \textbf{[S8.4]}$$

The initial age moments $\bar{\mathbf{a}}^{0,1}$ and $\bar{\mathbf{a}}^{0,2}$ can be obtained using the equilibrium formula

$$(-1)^n\,n!\,\left(\mathrm{X}^0\right)^{-1}\left(\mathrm{B}^0\right)^{-n}\mathbf{x^0},\quad n=1,2, \qquad \textbf{[S8.5]}$$

from ref. 12. Then $m_1 := \bar{a}_1^1(t)$ is the mean age of the atmospheric carbon at time t and $m_2 := \bar{a}_1^2(t)$ is its second moment. The SD can be computed as the square root of $m_2 - m_1^2$ from standard probability theory.

The trajectory of the age median of atmospheric carbon can be computed by solving Eq. **21** for $q=0.5$ and $i=1$. To that end, the cumulative compartment age distribution **P** can be obtained by Eq. **12** together with

$$\mathbf{P^0}(a) = \left(\mathrm{X}^0\right)^{-1}\left(\mathrm{B}^0\right)^{-1}\left(e^{a\,\mathrm{B}^0}-\mathrm{I}\right)\mathbf{u}(t_0),\quad a\geq 0, \qquad \textbf{[S8.6]}$$

where I is the three-dimensional identity matrix. To obtain Eq. **S8.6**, we need only to integrate Eq. **S8.2**. The initial age median ξ_1^0 of the atmospheric carbon at time t_0 needs to be approximated by a nonlinear optimization algorithm such that $P_1^0(\xi_1^0) = 0.5\,x_1^0$.

C. Forward Transit Time of Fossil Fuel-Derived Carbon. To compute the density of the forward transit time of fossil fuel-derived carbon, we simply change the input vector to $\mathbf{u}(t) = (u_A(t),0,0)^T$ and apply Eq. **27**. By using the new input vector, we consider the subsystem of only fossil fuel-derived carbon. We can treat this subsystem separately by means of the linear system that we derived by plugging the numerical solution into the nonlinear system.

Quantiles q, such as the median ($q=0.5$), for the forward transit time at arrival time t_a need to be computed by nonlinear optimization algorithms. To that end, $P_{\mathrm{FTT}}(\xi,t_a) = q\,\|\mathbf{u}(t_a)\|$ must be solved for ξ, where

$$P_{\mathrm{FTT}}(a,t_a) = \|\mathbf{u}(t_a)\| - \|\Phi(t_a+a,t_a)\,\mathbf{u}(t_a)\|. \qquad \textbf{[S8.7]}$$

Then, $t_a+\xi$ is the time at which the share q of the total input $\|\mathbf{u}(t_a)\|$ from time t_a will have left the system.

Other Supporting Information Files

Dataset S1 (CSV)