AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE

10.1002/2015WR017027

Special Section:
The 50th Anniversary of Water Resources Research

Key Points:
- Draw attention to overlooked theory of age and von Foerster equation
- First treatment of age distribution with random external forcing
- General results for soil moisture and streamflow models

Correspondence to:
A. Porporato,
amilcare@duke.edu

Citation:
Porporato, A., and S. Calabrese (2015), On the probabilistic structure of water age, *Water Resour. Res.*, *51*, 3588–3600, doi:10.1002/2015WR017027.

Received 2 FEB 2015
Accepted 10 APR 2015
Accepted article online 23 APR 2015
Published online 15 MAY 2015

© 2015. American Geophysical Union. All Rights Reserved.

On the probabilistic structure of water age

Amilcare Porporato[1] and Salvatore Calabrese[1]

[1]Department of Civil and Environmental Engineering, Duke University, Durham, North Carolina, USA

Abstract The age distribution of water in hydrologic systems has received renewed interest recently, especially in relation to watershed response to rainfall inputs. The purpose of this contribution is first to draw attention to existing theories of age distributions in population dynamics, fluid mechanics and stochastic groundwater, and in particular to the McKendrick-von Foerster equation and its generalizations and solutions. A second and more important goal is to clarify that, when hydrologic fluxes are modeled by means of time-varying stochastic processes, the age distributions must themselves be treated as random functions. Once their probabilistic structure is obtained, it can be used to characterize the variability of age distributions in real systems and thus help quantify the inherent uncertainty in the field determination of water age. We illustrate these concepts with reference to a stochastic storage model, which has been used as a minimalist model of soil moisture and streamflow dynamics.

1. Introduction

The description of the ages of a certain quantity (such as water, chemical compounds or individuals of a population) is fundamental in many fields. These include the characterization of hydrologic response of watersheds [*McGuire and McDonnell*, 2006], ecohydrological systems [*Botter et al.,* 2011], groundwater [*Cvetkovic and Dagan*, 1994; *Ginn*, 1999] and other geophysical systems [*Eriksson*, 1971; *Goode*, 1996; *Delhez et al.*, 1999], the performance of chemical reactors [*Nauman*, 1969, 2008], the structure of populations [*Kot*, 2001; *Murray*, 2002; *Ginn and Loge*, 2007] and the safety of manmade systems [*Cox*, 1962]. Owing to this wide interest, the literature on the subject is vast, and the previous references can only be considered as a starting point for more in-depth studies. In the field of watershed hydrology, this topic has witnessed a renewed interest in the recent years, with several works dealing with both theoretical and experimental aspects of age distribution of water in soils and watersheds (see e.g., the recent debate article by *McDonnell and Beven* [2014] and references therein).

The first goal of this paper is to point to the existence of some well-known literature in population dynamics, fluid mechanics and stochastic processes, in particular to the McKendrick-von Foerster (MKVF) equation [*M'Kendrick*, 1925; *von Foerster*, 1959; *Kot*, 2001; *Murray*, 2002], which is not commonly referenced in the hydrologic literature (indeed this was done only in one reference [*Ginn and Loge*, 2007]). It thus seems to be useful to briefly review these theories and place them under a common framework so that these results can be easily used. In particular, the link between the MKVF equation and its spatially explicit version, first proposed by *Ginn* [1999] [see also *Cvetkovic and Dagan*, 1994; *Ginn and Loge*, 2007; *Benettin et al.*, 2013], is important to emphasize the internal source of age variability introduced by spatial integration over a finite control volume.

The second goal is to address the role of external (random) hydrologic variability on the resulting age distributions. In fact, whenever hydrologic fluxes are treated stochastically on account of their unpredictable variability, the age distributions too become random functions and must be treated probabilistically. Thus, while the solution of the MKVF equation still offers the appropriate description of single realizations of such distributions (being suitable for any time varying conditions), their complete characterization requires considering the full ensemble of realizations of the underlying stochastic forcing. This gives rise to a probabilistic description of age distributions, which allows one to quantify their variability among realizations, thereby providing a starting point to quantify the uncertainty in their experimental determination. As far as we know this second aspect has not yet been explicitly addressed in the literature.

The paper is organized as follows. In section 2, we review the MKVF equation and connect it to the recent contributions in the hydrological literature. Section 3 illustrates the solutions of the MKVF equation for a

time varying system with step-change input as well as a steady state system with age-dependent output. Section 4 discusses the probabilistic aspect of age distributions, when inputs are modeled as stochastic processes, and section 5 presents an application to a linear stochastic storage system for which the probability density function (PDF) of the mean age is discussed. The conclusions briefly recapitulate the content of the paper and outline further open problems, in particular the role of nonlinearities and variability at different time scales.

2. Age Distribution Theory

Consider a control volume Ω with control surface Σ in which a substance, non necessarily conservative, is advected by the fluid velocity and diffused by molecular motion. To describe the age distribution it is useful to introduce the age-specific density or concentration, $c(\mathbf{x}, t, \tau)$, describing the quantity per volume at a point $\mathbf{x}$ and time t of a given substance having age τ. Note that typically c has units of mass per volume per age but, depending on the quantity under exam, it could also refer to age specific molar density or to age specific number of individuals/particles. The continuum hypothesis will be assumed to be valid.

A balance equation for c was first proposed by *Ginn* [1999] (see also *Delhez et al*. [1999] for a similar equation in the context of ocean dynamics)

$$\frac{\partial c}{\partial t}+\frac{\partial c}{\partial \tau}=-\nabla\cdot(\mathbf{v}c-D\nabla c)+\sigma, \tag{1}$$

where $\nabla=\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z}\right)$, $\mathbf{v}$ is the velocity field, σ is the net source term (due e.g., to chemical reactions or birth/death processes). The initial condition of the previous equation is the age distribution $c(\mathbf{x}, 0, \tau)$ for $\mathbf{x} \in \Omega$ and boundary conditions $c(\mathbf{x}, t, \tau)$ for $\mathbf{x} \in \Sigma$. We will make the assumption that age is counted from the time at which the quantity in consideration has entered the control volume (i.e., the mass is introduced in the control volume with zero age).

It is important to observe that the previous equation provides a purely kinematic description of age, in which the advective velocity is assumed to be known or given by coupled equations of motion. In the case of a hydrological systems, however, this picture typically requires an unattainable amount of information [*McDonnell and Beven*, 2014], so that a spatially lumped description over the entire control volume becomes necessary, along with auxiliary assumptions to surrogate for the lack of knowledge of the velocity field. Another interesting point regards the fact that both the velocity vector and the net source/sink term in (1) may also be a function of age, especially in the case of active populations (bacterial, animals, people), or in hydrological systems where for example transpiration may select water in specific ranges of ages depending on the overall level of moisture, because of soil-pore heterogeneities.

In case of lack of detailed kinematic description, as is the case of hydrological systems, it is then useful to integrate over the control volume to achieve a spatially lumped formulation for the age distribution function,

$$n(t,\tau)=\int_{\Omega} c(t,\tau,\mathbf{x})d\Omega, \tag{2}$$

which quantifies the amount of the substance under exam at a given time, in units of mass per age for the entire control volume. When the local equation (1) is integrated over the control volume using the Gauss theorem [*Ginn and Loge*, 2007; *Benettin et al*., 2013], the MKVF equation [*M'Kendrick*, 1925; *von Foerster*, 1959] is readily obtained

$$\frac{\partial n}{\partial t}+\frac{\partial n}{\partial \tau}=-\mu n, \tag{3}$$

with initial condition $n(t=0,\tau)=\beta(\tau)$ for the age distribution and boundary condition $n(t,\tau=0)=\alpha(t)$, also called the birth rate. The so-called mortality rate, or loss term, $\mu(t,\tau)$, may be subdivided into different terms according to the different pathways out of the system (e.g., in surface hydrology, evapotranspiration, percolation, runoff, etc.) and can be considered a nonlinear function the total storage w, to be defined shortly, as in *Gurtin and MacCamy* [1974, 1979] and subsequent papers. A direct derivation of the MKVF equation, without going through equation (1), can be found in several books and articles [*Trucco*, 1965; *Keyfitz and Keyfitz*, 1997; *Kot*, 2001; *Murray*, 2002].

Before reviewing the solution of (3), it is useful to connect it to the familiar form of the mass balance equation for a control volume. Indicating as

$$w(t)=\int_0^\infty n(t,\tau)d\tau, \tag{4}$$

the total mass in Ω at time t, integration of (3) gives

$$\frac{dw}{dt}=\imath-o, \tag{5}$$

where the input and output terms are respectively

$$\imath(t)=n(t,0)=\alpha(t) \text{ and } o(t)=\int_0^\infty \mu n d\tau. \tag{6}$$

It may also be useful to define the normalized distribution of ages,

$$\phi(t,\tau)=\frac{n(t,\tau)}{w(t)}, \tag{7}$$

which can be interpreted as a probability density function (PDF) of the ages when the quantity under exam is sampled at random in Ω. It is important to keep in mind that the age variability described by this distribution here results from the spatially lumped description and reflects the possible complex streamline structure (e.g., due to turbulence or heterogeneity of the porous medium) within Ω but does not stem from the external temporal hydrologic variability, which will be introduced later (see section 4).

The mean age is defined as

$$a(t)=\int_0^\infty \tau\phi(t,\tau)d\tau=\frac{\int_0^\infty \tau n(t,\tau)d\tau}{w(t)}=\frac{m(t)}{w(t)} \tag{8}$$

and similarly for higher order moments. Solving (7) for n and substituting into (3) one obtains the equation of *Botter et al*. [2011], also used by *Harman* [2014]. While being of course an exact consequence of the MKVF equation, this resulting equation is in reality less convenient for solution (note that *van der Velde et al*. [2012] instead write directly the MKVF equation, although they attribute it to *Botter et al*. [2011]). In this paper we will basically only refer to the age distribution $n(t,\tau)$, its normalized pdf version $\phi(t,\tau)$ and its first moment $a(t)$. Other important related quantities, such as the residence time and the survival time distributions [*Nauman*, 1969; *Cox*, 1962; *Niemi*, 1977], along with their link to the the survival equation and renewal theory [*Cox*, 1962], will be discussed elsewhere. We will only recall, as already noted by *von Foerster* [1959] [see also *Manzoni et al*., 2009], that in steady state (we use the suffix * to emphasize this fact) the age distribution, suitably normalized by the constant input, gives the survivor function of the system [*Cox*, 1962; *Daly and Porporato*, 2006a],

$$\varphi^*(\tau)=\frac{n^*(\tau)}{\imath}. \tag{9}$$

In turn, the negative of its derivative with respect to age directly provides the distribution of life span or transit times in the system,

$$\psi^*(\tau)=-\frac{d\varphi^*(\tau)}{d\tau}. \tag{10}$$

It is important to emphasize that these quantities are in general different for time-varying conditions and that here we will mostly focus on the age distributions.

Returning to the MKVF equation, its solution can be obtained using the method of characteristics, which is defined by $\frac{d\tau}{dt}=1$ [*Trucco*, 1965; *Murray*, 2002], expressing the obvious fact that age increases at the same rate as time. The solution is [*Trucco*, 1965]

$$n(t,\tau)=\begin{cases}\beta(\tau-t)\,e^{-\int_0^t \mu(u,\tau-t+u)du} & t<\tau\\ \alpha(t-\tau)\,e^{-\int_0^\tau \mu(t-\tau+u,u)du} & t>\tau.\end{cases} \tag{11}$$

We specifically refer to the case in which the "mortality" term is only a function of age and time, $\mu(t,\tau)$, but not of the total quantity w (this is the nonlinear case dealt with by *Gurtin and MacCamy* [1974, 1979]). When the input is in form of instantaneous pulses (i.e., Dirac deltas), in which case the age distribution is also a singular function, it may be convenient to introduce the auxiliary functions

$$u(t,\tau)=\int_\tau^\infty n(t,\tau')d\tau' \tag{12}$$

and

$$v(t,\tau)=\int_{-\infty}^t n(t',\tau)dt', \tag{13}$$

being $u(t,\tau)$ the total amount of the quantity having age greater than τ at time t (when the integral starts from 0, it is equal to $w(t)$, see (4)), and $v(t,\tau)$ the cumulative amount of substance, from the beginning to time t, having had age equal to τ.

Finally, a useful hierarchy of equations for the moments of n can be also obtained, similarly to what was done in *Rotenberg* [1972], *Varni and Carrera* [1998], and *Duffy* [2010] for the point equation (1). Focusing in particular on the mean age, we can first multiply (3) by τ and then integrate with respect to age to write

$$\frac{dm}{dt}=w-\theta, \tag{14}$$

where $\theta=\int_0^\infty \tau\mu n d\tau$ (this could be interpreted as the first moment of the age distribution at exit or death). An interesting equation for the mean age $a=m/w$ can then be obtained by differentiating the definition and using the previous equations (5) and (14),

$$\frac{da}{dt}=1-a\left(\frac{\theta}{m}+\frac{\iota}{w}-\frac{o}{w}\right). \tag{15}$$

In the case of age-independent mortality, say $\mu=\eta=const.$, then

$$\frac{da}{dt}=1-a\frac{\iota}{w}. \tag{16}$$

While obviously not in closed form, being coupled to (5), this equation is however quite instructive for it shows how with age-independent outputs the system simply ages with time in the absence of input, which in turn acts to "dilute" the overall age of the system, whereas the term w brings the whole history of the input into the dynamics of the mean age.

3. Examples

In this section, we present three simple applications to illustrate the previous theory. The first one emphasizes the effect of time varying input as a prelude to its randomization in section 5, while the second and third applications focus on the role of age-dependence in the mortality term.

3.1. Time-Varying Input

We first consider the case of constant mortality rate η, which corresponds to the classical linear system, extensively used in hydrology [e.g., *Singh*, 1992]. Starting from an initial condition $n(0,\tau)=n_0(\tau)$, the system proceeds with constant input ι_0, until at time t_1 the input is suddenly switched to a new value ι_1. From equation (11), the transient solution is

$$n(t,\tau)=\begin{cases}n_0(\tau-t)\,e^{-\eta t} & t<\tau\\ \iota_0(t-\tau)\,e^{-\eta\tau} & t_1>t>\tau\\ \iota_1(t-\tau)\,e^{-\eta\tau} & t>t_1\end{cases} \tag{17}$$

which, for $t\to\infty$, tends to an exponential steady state solution. The behavior of the various solutions is shown in Figure 1 for a uniform initial condition over the age range $(0,w_0/n_0)$. Note the steady state value

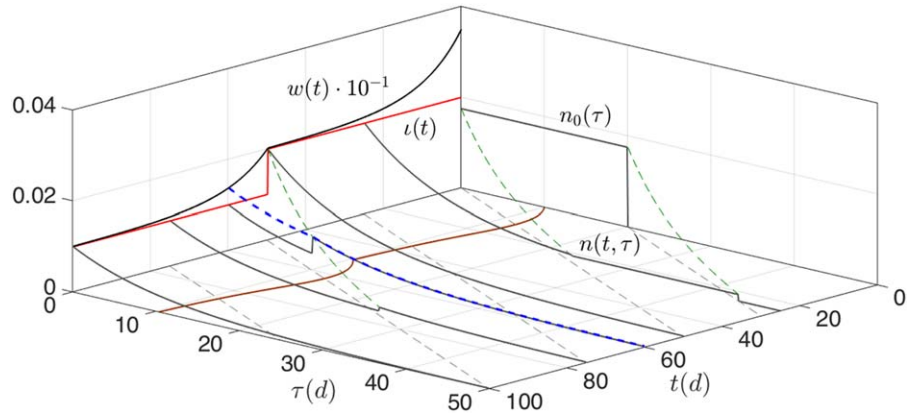

Figure 1. Time evolution of the soil water age distribution (black) driven by step change of the input (red). The total storage $w(t)$ is plotted in blue on the vertical plane on the left, while the mean age $a(t)$ is on the horizontal plane in brown. The dashed blue line represents the variable $u(t,\tau)$ for a given t and the dotted green lines show the decays along the characteristic curves (dashed straight lines on the horizontal plane). $\iota_1 = 2\iota_2 = 2 \cdot 10^{-2}$ [w]d^{-1} and $\eta = 0.1$ d^{-1}.

of the mean age, equal to $1/\eta$, independent of the input rate, which for this linear case is also the mean life span (as follows from (10) and the fact that both ϕ^* and φ^* are exponential distributions). However, for the purposed of this paper, it is important to emphasize that in the time-varying case these are in general different distributions.

3.2. Age-Dependent Loss

This application regards a system that reaches steady state under constant input but age-dependent loss term. The loss term is chosen in the form of a power law,

$$\mu(\tau) = \left(\frac{\tau}{1+c}\right)^c, \tag{18}$$

($c > -1$), representing for example the effects of an age-dependent plant water uptake in soils, where the parameters c could be related to the pore size distribution and tortuosity of the flow paths, as in power-law soils [*Brutsaert*, 2005].

Figure 2 shows the temporal evolution of the various terms on the right hand side of equation (15) for this case. At steady state, the age distribution is

$$n^*(\tau) = \iota \exp\left(-\left(\frac{\tau}{1+c}\right)^{1+c}\right), \tag{19}$$

whose integral with respect to τ from 0 to ∞ gives

$$w^* = \iota\, \Gamma\left(\frac{1}{1+c}\right), \tag{20}$$

while the mean age is,

$$a^* = (1+c)\frac{\Gamma\left(\frac{2}{1+c}\right)}{\Gamma\left(\frac{1}{1+c}\right)}. \tag{21}$$

Finally, using (10) one obtains

$$\psi^*(\tau) = \left(\frac{\tau}{1+c}\right)^c \exp\left(-\left(\frac{\tau}{1+c}\right)^{1+c}\right), \tag{22}$$

which is a Weibull distribution with scale parameter and shape parameter both equal to $1+c$. The mean of ψ^*, the mean residence time, is $\Gamma(\frac{1}{1+c})$, always greater than a^*, as expected in this case [*Bolin and Rodhe*, 1973; *Björkström*, 1978], and the variance is $2(1+c)\Gamma(\frac{2}{1+c}) - (\Gamma(\frac{1}{1+c}))^2$. The Weibull distribution is frequently used in renewal theory [*Cox*, 1962; *Daly and Porporato*, 2006a] and, as well known, is one of the Extreme Value distributions [*Kottegoda and Rosso*, 1997; *Sornette*, 2006].

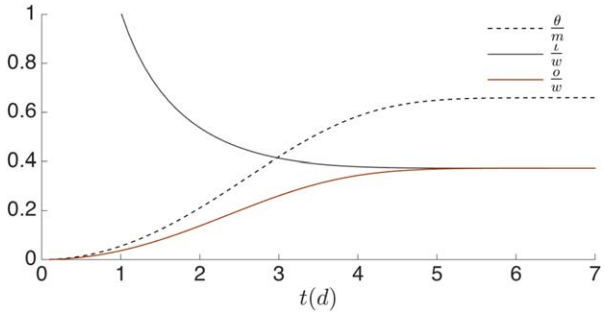

Figure 2. Time evolution toward steady state of the terms $\frac{\theta}{m}$, $\frac{\iota}{w}$ and $\frac{o}{w}$ from equation (15) for the case of age-dependent loss.

3.3. Heavy Tails in Age Distributions

In this last example, we show how a particular form of age-dependence in the loss function may be related to the observed different time scales in the hydrologic response of watersheds and give rise to heavy tails in the water age distributions [see e.g., *McDonnell and Beven*, 2014; *Kirchner et al.*, 2000; *Rodhe et al.*, 1996]. If one chooses a hyperbolic behavior of the loss function with age, such that there is a strong preference for young water to be lost, coupled with a very slow decay of the loss function at large ages,

$$\mu(\tau)=\frac{1}{\sigma\left(\frac{\epsilon(\tau-\mu)}{\sigma}+1\right)}, \tag{23}$$

the survivor function $\varphi^*(\tau)$ and life span distribution $\psi^*(\tau)$ are then,

$$\varphi^*(\tau)=\left(\frac{\epsilon(\tau-\mu)}{\sigma}+1\right)^{-1/\epsilon} \tag{24}$$

$$\psi^*(\tau)=\frac{1}{\sigma}\left(\frac{\epsilon(\tau-\mu)}{\sigma}+1\right)^{-\left(\frac{1}{\epsilon}+1\right)}. \tag{25}$$

The function $\psi^*(\tau)$ is a Generalized Pareto distribution [*Pareto*, 1964; *Kottegoda and Rosso*, 1997], with mean and variance respectively equal to $\mu+\frac{\sigma}{1-\epsilon}$ and $\frac{\sigma^2}{(1-\epsilon)^2(1-2\epsilon)}$. It belongs to the class of Heavy-tailed distributions. As shown in Figure 3b, both the survivor and transit time distributions are characterized by the presence of algebraic decays in the tail behavior (i.e., heavy tails). Differently from the previous example, the loss function here is such that the mean age is greater than mean transit time [*Bolin and Rodhe*, 1973].

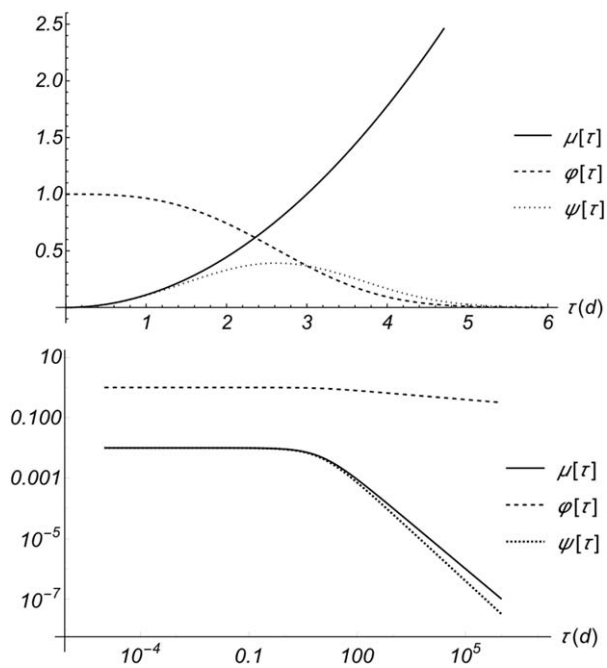

Figure 3. Steady state solutions for the survivor functions $\varphi^*(\tau)$, the transit time distributions $\psi^*(\tau)$, shown along with the loss functions of (a) equation (18) with parameter $c=2$ and (b) equation (23), in log-log scale, with parameters $\mu=0$, $\sigma=100$, and $\epsilon=10$.

4. Effect of External Stochasticity

From the previous discussion it is evident that the age distribution presented before reflects the integration of a flow field over the control volume. This variability can thus be called "internal." In the case of hydrological systems, however, the effect of random "external" forcing should also be taken into account, since it may be at least as important as the internal one. In what follows we will limit our analysis to the more important case of stochastic inputs, while considering outputs as deterministic. As discussed by *Daly and Porporato* [2006b], the rainfall-induced variability in fact tends to dominate over the variability in the outputs (e.g., evapotranspiration, percolation, and runoff).

From the mathematical point of view, the presence of a random input makes (5) a

stochastic differential equation (SDE). While the MKVF equation is still perfectly suitable, in that a single realization of the random input is just a special form of time varying input entering through the boundary conditions at age zero, it is important to keep in mind that the age distribution too becomes a random function, changing for each realization of the stochastic input. Thus, to move beyond a single realization and obtain a complete probabilistic description, one needs to consider the ensemble of all possible realizations of the random input and introduce related probability distributions.

To emphasize the randomness resulting from the stochastic forcing, we will use capital letters for random variables, as usual in statistics. Accordingly, one can define the random function $N(t,\tau)$, and consider its ensemble average $\langle N(t,\tau)\rangle$, whose equation follows directly from the MKVF (when μ does not depend on N or W),

$$\frac{\partial\langle N\rangle}{\partial t}+\frac{\partial\langle N\rangle}{\partial\tau}=-\mu\langle N\rangle, \tag{26}$$

driven by the ensemble average of the random input, entering as the $\tau=0$ boundary condition. However, there is no need to resort to the previous equation, because the general solution (11) is still valid, as long as one considers the randomly varying input, $I(t)$, as boundary condition,

$$N(t,\tau)=\begin{cases}\beta(\tau-t)\,e^{-\int_0^t\mu(u,\tau-t+u)du} & t<\tau\\ I(t-\tau)\,e^{-\int_0^\tau\mu(t-\tau+u,u)du} & t>\tau\end{cases}, \tag{27}$$

where the initial age distribution can be considered as an assigned random function with its own probabilistic structure. This solution is interesting for it shows that for a given time and age, the age density is just the rescaled in time version of the random input. Thus for the one-time, one-age PDF of N, which is formally connected to N by ensemble average of its fine grained PDF [*Stratonovich*, 1967; *Pope*, 2000], $p_N(n;t,\tau)=\langle\delta(N(t,\tau)-n)\rangle$, and represents the infinitesimal probability to observe an amount between n and $n+dn$ with age τ at time t, one can write (for $t>\tau$)

$$p_N(n;t,\tau)=e^{\int_0^t\mu(u,\tau-t+u)du}p_I(ne^{\int_0^t\mu(u,\tau-t+u)du};t-\tau), \tag{28}$$

where $p_I(\imath;t)$ is the one time PDF of the input.

One can also define $W=\int_0^\infty N(t,\tau)$ and write, thanks to the linear operations of averaging,

$$\langle W(t)\rangle=\int\langle N(t,\tau)\rangle d\tau, \tag{29}$$

and similarly for the auxiliary variables $U(t,\tau)$ and $V(t,\tau)$, see (12) and (13).

The mean age,

$$A(t)=\frac{M(t)}{W(t)}=\frac{\int_0^\infty\tau N(t,\tau)d\tau}{\int_0^\infty N(t,\tau)d\tau}, \tag{30}$$

also acquires random variability, governed by the SDE version of equation (15). A general and relatively simple solution for the harmonic mean of the mean age A can be obtained by dividing by M the SDE equivalent of (14), i.e.,

$$\frac{d\ln M}{dt}=\frac{W}{M}-\frac{\Theta}{M} \tag{31}$$

where $\Theta=\int_0^\infty\tau\mu N d\tau$. In conditions of statistical steady state, taking ensemble average, this gives the exact result

$$\left\langle\frac{1}{A}\right\rangle^{-1}=\left\langle\frac{\Theta}{M}\right\rangle^{-1}. \tag{32}$$

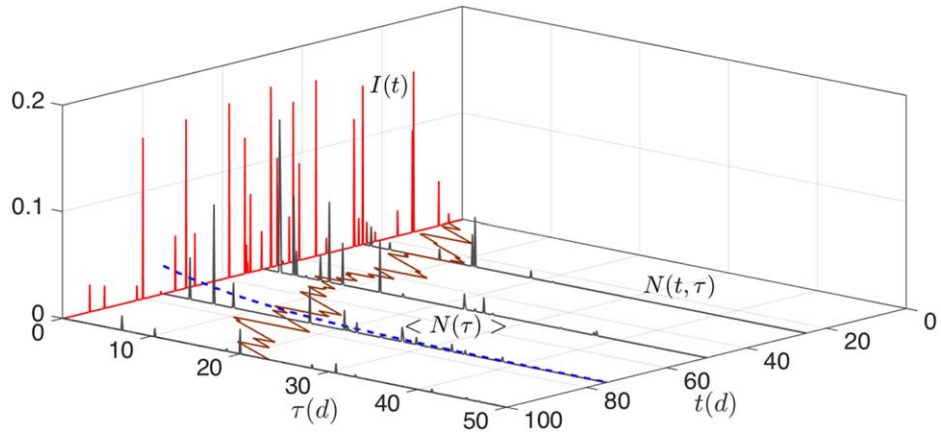

Figure 4. Time evolution of soil water age distribution (solid gray line) along with the marked Poisson process inputs (red line in the vertical plane on the left) with mean depth $1/\gamma$=2 mm and mean frequency λ=0.4 d^{-1}. The ensemble average age distribution $\langle N \rangle$ is plotted at t = 75 d (dashed blue line), while the mean age is plotted in brown on the horizontal plane. Note that for graphical purposes I and N are plotted using the size of the marks, while in theory they are singular functions made of infinitely high spikes (Dirac deltas).

For a linear system $\mu=\eta=const.$, then immediately

$$\left\langle \frac{1}{A} \right\rangle^{-1} = \frac{1}{\eta}. \tag{33}$$

5. Stochastic Linear Storage

We apply the previous concepts to the stochastic linear water balance equation

$$\frac{dW}{dt} = I - \eta W, \tag{34}$$

where η is a constant and the input I is the formal time derivative of a marked Poisson process with rate (frequency of occurrence) λ and exponentially distributed marks with mean $1/\gamma$ (a realization of the input is shown in the left vertical plane of Figure 4). This SDE has been used as a minimalist model of both soil moisture, in conditions when evapotranspiration losses dominate over percolation and runoff [*Isham et al.*, 2005; *Viola et al.*, 2008], and streamflow dynamics [*Weiss*, 1977; *Botter et al.*, 2007]. An example of a realization of the process $W(t)$, a sequence of exponential jumps and exponential decays of exponential duration, is shown in the left vertical plane of Figure 5. The one time PDF, $p_W(w;t)$, is a hypergeometric function, which becomes a gamma distribution in stochastic steady state [*Viola et al.*, 2008]. It may be useful to warn the reader not to confuse the deterministic steady state condition discussed in sections 2 and 3 with the statistical steady state discussed here, where time varying conditions induced by the external forcing are always present.

Because of the intermittent nature of the input, when considering only one realization, the water in the system only has ages corresponding to the times of previous input occurrences, while there will be no water for all other ages. As a result, in the present case, $N(t,\tau)$ is a singular function, being the exponentially modulated version of the time derivative of the marked Poisson input (i.e., a sequence of instantaneous infinite spikes). This is shown in Figure 5. A convenient way to avoid handling such singular functions is to refer to their integrated auxiliary variables, U and V, also shown in Figure 5. Note that $U(t,0)=W(t)$, so that its distribution is expected to be a hypergeometric function, at least for $\tau=0$, while $V(t,\tau)$ is the exponentially rescaled version of the total rainfall up to that time (whose distribution is known to be a Bessel function). These distributions of $p_U(u;t,\tau)$ and $p_V(v;t,\tau)$ are derived in detail in Appendix A. An example of their shapes is shown in Figure 6, along with some realizations.

As for the mean age, $A=M/W$, we already know from (33) that its harmonic mean is equal to $1/\eta$. However, to obtain its full distribution one should first derive the joint distribution of W and M, $p_{WM}(w,m;t)$, and then use the formula for the quotient of random variables [*Kottegoda and Rosso*, 1997],

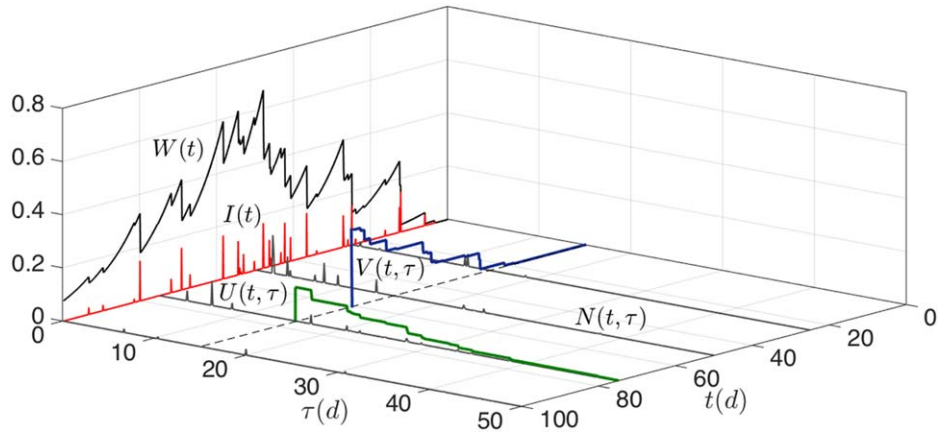

Figure 5. For the same realization of Figure 4, $W(t)$ is plotted on the left vertical plane (solid black line). The variables $U(t)$, and $V(t)$ (solid green and brown lines, respectively) are also shown. Parameters as in Figure 4.

$$P_A(a;t)=\int_{-\infty}^{\infty} z p_{MW}(az,z;t)dz. \tag{35}$$

Thus, starting from the coupled SDEs

$$\begin{cases} \dfrac{dW}{dt}=I-\eta W \\ \dfrac{dM}{dt}=W-\eta M, \end{cases} \tag{36}$$

the Chapman-Kolmogorov equation for the joint PDF can be written as in *Botter et al.* [2008],

$$\frac{\partial p_{WM}}{\partial t}=\frac{\partial}{\partial w}(w\eta p_{WM})+\frac{\partial}{\partial M}((\eta m-w)p_{WM})-\lambda p_{WM}+\lambda\gamma\int_0^w p_{WM}(x,m;t)e^{-\gamma(w-x)}dx. \tag{37}$$

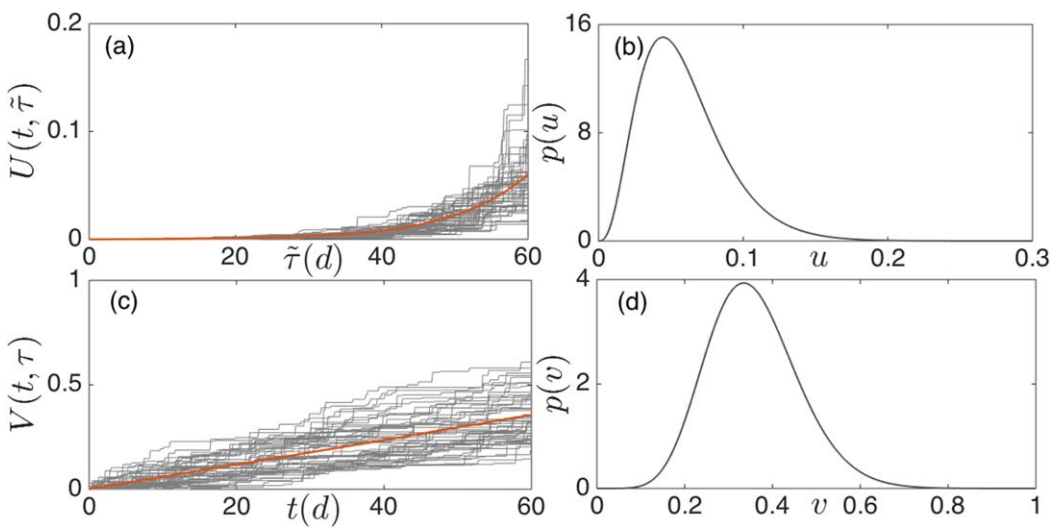

Figure 6. Traces of realizations of U (a) as a function of $\tilde{\tau}$ as defined in Appendix A, equation (A1) and V (c) as a function of time t. Probability density functions $p_U(u;t,\tau)$ (b) for $t=75$ d and $\tilde{\tau}=60$ d, and $p_V(v;t,\tau)$ (d) for $t=60$ d and $\tau=15$ d.

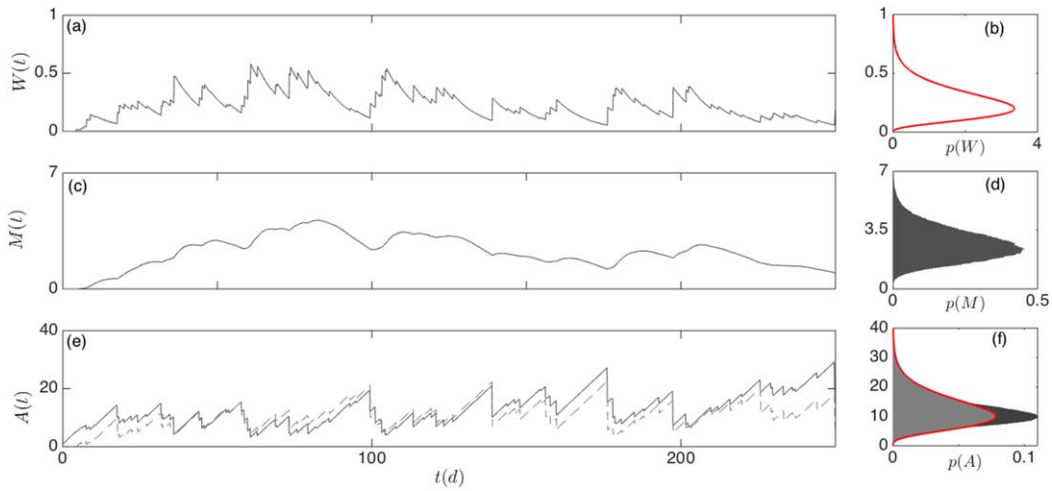

Figure 7. Time evolution of the variables (a) $W(t)$, (c) $M(t)$ and (e) $A(t)$. In Figure 7e, solid and dashed lines refer to the solution of equation (38) respectively with and without the assumption $W \approx \langle W \rangle = \lambda/(\gamma\eta)$. (b) pdf of $W(t)$. (d) Histogram of $M(t)$. (f) Analytical pdf and histogram of $A(t)$ (with $W \approx \langle W \rangle = \lambda/(\gamma\eta)$) (solid red and light gray), and histogram for the real solution $A(t)$ (dark gray).

While the moments of the joint distribution can be readily computed through the moment generating function (Laplace transform), see e.g., *Botter et al.* [2008], the full joint distribution is unfortunately very difficult to obtain. An alternative, approximate route will be followed here. Starting from the equation for A, see (16), and by making the approximation $W \approx \langle W \rangle$, so that the equation becomes closed and Markovian,

$$\frac{dA}{dt} = 1 - \frac{A}{\langle W \rangle} \circ I, \tag{38}$$

where the symbol $\circ$ underlines the fact that in this SDE the multiplication of the random variable A by the white noise I should be interpreted in the sense of Stratonovich [*Gardiner*, 2009; *Suweis et al.*, 2011] in order to preserve the ordinary rules of calculus which are used to derive the equation for A as a derivative of a quotient, $A = M/W$. Then, being $\langle W \rangle = \lambda/(\gamma\eta)$ and using the transformation $B = \ln A$, an additive noise SDE for B can be obtained as

$$\frac{dB}{dt} = e^{-B} - \frac{\gamma}{\lambda} I, \tag{39}$$

whose solution in steady state can be found as in *Suweis et al.* [2010]. Reverting to the PDF of A using the random-variable function theorem, results in the approximate PDF of A in the form of a gamma distribution,

$$p_A(a) = \frac{\lambda^{\frac{\lambda}{\eta}+1}}{\Gamma(\frac{\lambda}{\eta}+1)} a^{\frac{\lambda}{\eta}} e^{-\lambda a}, \tag{40}$$

with mean

$$\langle A \rangle = \frac{1}{\eta} + \frac{1}{\lambda}, \tag{41}$$

variance

$$\mathrm{var}_A = \frac{1}{\lambda}\left(\frac{1}{\eta} + \frac{1}{\lambda}\right), \tag{42}$$

and mode

$$\mathrm{mode}_A = \frac{1}{\eta}. \tag{43}$$

For a gamma distribution, the mode is also the harmonic mean, which, interestingly, is an exact result in agreement with (33). Figure 7 compares the approximate solution with the one obtained from numerical simulations of the full system (36). The approximate PDF captures well the shape and location of the simulated PDF, however the approximate one tends to have weaker mode and larger spread. More detailed comparisons of the validity of the used approximation will be presented in subsequent contributions. For this paper, the approximate PDF of the mean age already serves the purpose of showing how the (external) random variability in the input propagates into the probabilistic properties of age. As seen in Figure 7, even the mean water age can vary considerably from sample to sample by virtue of the input variability. This type of variability could thus be expected to significantly impact the year-to-year determination of water age in a hydrologic system. One should also mention that the external stochasticity in the input may be responsible for a heavier tail behavior of the age distribution (see distribution of mean age in Figure 7), compared to the case with only internal variability. This effect is well known in compound distributions [*Benjamin and Cornell*, 1970] and has recently been advocated as a mechanisms for nonextensive thermodynamics behavior in statistical mechanics [*Beck and Cohen*, 2003] and for increased variability of ecohydrologic variables [*Porporato et al*., 2006].

6. Conclusions

While revisiting the general solution of the MKVF equation within the context of the theory of age distributions, we have emphasized that such distributions by their nature stem from integrating the internal structure of flow pathways within hydrologic control volumes [*Ginn and Loge*, 2007]. Such "internal" variability should be contrasted with another form of "external" stochasticity, originating from the randomness inherent in input and output terms of hydrologic systems. The MKVF equation and the stochastic differential equations that can be derived from it suggest a framework to include these external stochastic effects and quantify their contribution to the probabilistic structure of water age. Some results regarding the statistical properties of mean age have been described in sections 4 and 5. Even for the age independent loss function of sections 3.1 and 5, the temporal variability of external input induces non trivial effects on the age distribution.

Several issues remain to be explored along these lines, including the effects of age-dependence and nonlinearities in the loss function, as well as their random variability at different time scales (e.g., daily, seasonal and interannual) and the link to probabilistic renewal theory. This notwithstanding, we hope that this work could represent a useful step toward improving the characterization water age distributions in hydrologic and geophysical systems.

Appendix A

We derive here the PDFs of the auxiliary variables $U(t,\tau)=\int_{\tau}^{\infty} N(t,\tau')d\tau'$ and $V(t,\tau)=\int_{-\infty}^{t} N(t',\tau)dt'$ for the linear stochastic storage of section 5. As evident from Figure 6 such variables are the cumulative processes of a marked Poisson process. In particular, introducing for convenience an auxiliary time variable,

$$\tilde{\tau}=t-\tau \tag{A1}$$

to be used instead of τ, then

$$U(t,\tilde{\tau})=-\int_{\infty}^{t-\tilde{\tau}} n(t,\tilde{\tau}')d\tilde{\tau}'. \tag{A2}$$

This is the integral of the formal time derivative of a marked Poisson process with rate λ and jumps with parameter $g(\tilde{\tau})=\gamma e^{\eta(t-\tilde{\tau})}$. The corresponding master equation for $p_U(u;t,\tilde{\tau})$ is [*Cox and Miller*, 1977]

$$\frac{\partial p_U}{\partial \tilde{\tau}}=-\lambda p_U+\lambda g(\tilde{\tau})\int_0^u p_U\, e^{-g(\tilde{\tau})(u-x)}dx, \tag{A3}$$

with initial condition $p_U(u;t,0)=\delta(u-u_0)$, with $\delta(\cdot)$ the Dirac delta function and $u_0=w(t=0)e^{-\eta t}$. By taking the Laplace transform with respect to $u\rightarrow k$, $p_U(u;t,\tilde{\tau})\rightarrow p_U^{\dagger}(k;t,\tilde{\tau})$, one obtains

$$\frac{dp_U^{\dagger}}{d\tilde{\tau}} = -\lambda p^{\dagger}\left(1-\frac{g(\tilde{\tau})}{g(\tilde{\tau})+k}\right), \tag{A4}$$

which can be solved by dividing by $p_U^{\dagger}$, integrating with respect to $\tilde{\tau}$ and then taking the exponential,

$$p_U^{\dagger}(k;t,\tilde{\tau}) = p_U^{\dagger}(k;t,0)e^{-\lambda\tilde{\tau}}\left(1-\frac{\gamma e^{\eta\tilde{\tau}}}{\gamma e^{\eta\tilde{\tau}}+k}\right). \tag{A5}$$

By substituting $y = ke^{-\eta\tilde{\tau}}$ in the previous expression one obtains the Laplace transform of a jump process with exponential distributed jumps and linear drift, whose inverse transform has already been given in *Tsurui and Osaki* [1976]; *Viola et al*. [2008]; *Daly and Porporato* [2010]. Performing then the transformation of random variables, $p(k) = p(y)\frac{dk}{dy}$, and knowing that $\tilde{\tau} = t-\tau$, one finally obtains the desired solution, plotted in Figure 6b,

$$p_U(u;t,\tau) = \vartheta(u-u_0)e^{\eta\tau}\left(e^{-\eta(t-\tau)}\delta(u-u_0) + e^{-\eta(t-\tau)-\gamma(u-u_0)}\frac{\lambda\gamma}{\eta}(e^{\eta(t-\tau)}-1)\,{}_1F_1\left((1-\frac{\lambda}{\eta},2,\gamma(u-u_0)(1-e^{\eta(t-\tau)})\right)\right), \tag{A6}$$

where $\vartheta(\cdot)$ is the heaviside function and ${}_1F_1(\cdot,\cdot)$ is the Kummer confluent hypergeometric function [*Abramowitz and Stegun*, 2012].

Similarly, the variable V represents the integral over time of the derivative of a marked Poisson process with the same rate λ and rescaled parameter for the jumps, $g(\tau) = \gamma e^{-\gamma\tau}$. Note that the latter is a now a constant with respect to the integration variable t and the initial condition is $p_V(v;0,\tau) = \delta(v-v_0)$ and $v_0 = w(t=0)e^{-\eta\tau}$. As a result, considering constant g in (A4), and integrating with respect to t,

$$p_V^{\dagger}(k;t,\tau) = e^{-v_0 k}e^{-\lambda t}e^{\frac{\lambda g(\tau)}{g(\tau)+k}t}. \tag{A7}$$

Antitransforming [see e.g., *Daly and Porporato*, 2010], the solution (see Figure 6d) is found as

$$p_V(v;t,\tau) = \vartheta(v-v_0)e^{-\eta t}e^{-g(\tau)(v-v_0)}\left(\sqrt{\frac{\lambda g(\tau)t}{v-v_0}}I_1(2\sqrt{\lambda g(\tau)(v-v_0)t}) + \delta(v-v_0)\right), \tag{A8}$$

where $I_1(\cdot)$ is the modified Bessel function of the first kind of order 1 [*Abramowitz and Stegun*, 2012].

Acknowledgments
This work was partially funded through the Agriculture and Food Research Initiative of the USDA National Institute of Food and Agriculture (2011-67003-30222); the National Science Foundation through grants CBET-1033467, EAR-1331846, FESD-1338694, and EAR-1316258; and by the U.S. Department of Energy (DOE) through the Office of Biological and Environmental Research (BER) Terrestrial Carbon Processes Program (DE-SC0006967). We gratefully acknowledge useful discussions with I. Rodriguez-Iturbe, E. Daly and C. Duffy during the preparation of this manuscript. This paper is a theoretical study and hence no data were used.

References

Abramowitz, M., and I. A. Stegun (2012), *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, Courier Dover Publ., N. Y.
Beck, C., and E. Cohen (2003), Superstatistics, *Physica A*, *322*, 267–275.
Benettin, P., A. Rinaldo, and G. Botter (2013), Kinematics of age mixing in advection-dispersion models, *Water Resour. Res.*, *49*, 8539–8551, doi:10.1002/2013WR014708.
Benjamin, J. R., and C. A. Cornell (1970), *Probability, Statistics, and Decision for Civil Engineers*, McGraw-Hill, N. Y.
Björkström, A. (1978), A note on the relation between average age and average transit time in natural reservoirs, *Tellus*, *30*(2), 185–188.
Bolin, B., and H. Rodhe (1973), A note on the concepts of age distribution and transit time in natural reservoirs, *Tellus*, *25*(1), 58–62.
Botter, G., A. Porporato, I. Rodriguez-Iturbe, and A. Rinaldo (2007), Basin-scale soil moisture dynamics and the probabilistic characterization of carrier hydrologic flows: Slow, leaching-prone components of the hydrologic response, *Water Resour. Res.*, *43*, W02417, doi:10.1029/2006WR005043.
Botter, G., E. Daly, A. Porporato, I. Rodriguez-Iturbe, and A. Rinaldo (2008), Probabilistic dynamics of soil nitrate: Coupling of ecohydrological and biogeochemical processes, *Water Resour. Res.*, *44*, W03416, doi:10.1029/2007WR006108.
Botter, G., E. Bertuzzo, and A. Rinaldo (2011), Catchment residence and travel time distributions: The master equation, *Geophys. Res. Lett.*, *38*, L11403, doi:10.1029/2011GL047666.
Brutsaert, W. (2005), *Hydrology: An Introduction*, Cambridge Univ. Press, Cambridge, U. K.
Cox, D. R. (1962), *Renewal Theory*, Methuen, London, U. K.
Cox, D. R., and H. D. Miller (1977), *The Theory of Stochastic Processes*, vol. 134, CRC Press, Boca Raton, Fla.
Cvetkovic, V., and G. Dagan (1994), Transport of kinetically sorbing solute by steady random velocity in heterogeneous porous formations, *J. Fluid Mech.*, *265*, 189–215.
Daly, E., and A. Porporato (2006a), State-dependent fire models and related renewal processes, *Phys. Rev. E*, *74*(4), 041112.
Daly, E., and A. Porporato (2006b), Impact of hydroclimatic fluctuations on the soil water balance, *Water Resour. Res.*, *42*, W06401, doi:10.1029/2005WR004606.
Daly, E., and A. Porporato (2010), Effect of different jump distributions on the dynamics of jump processes, *Phys. Rev. E*, *81*(6), 061133.
Delhez, E. J., J.-M. Campin, A. C. Hirst, and E. Deleersnijder (1999), Toward a general theory of the age in ocean modelling, *Ocean Modell.*, *1*(1), 17–27.
Duffy, C. J. (2010), Dynamical modelling of concentration–age–discharge in watersheds, *Hydrol. Processes*, *24*(12), 1711–1718.
Eriksson, E. (1971), Compartment models and reservoir theory, *Annu. Rev. Ecol. Syst.*, *2*, 67–84.
Gardiner, C. W. (2009), *Stochastic Methods: A Handbook for the Natural and Social Sciences*, Springer, Berlin.

Ginn, T. R. (1999), On the distribution of multicomponent mixtures over generalized exposure time in subsurface flow and reactive transport: Foundations, and formulations for groundwater age, chemical heterogeneity, and biodegradation, *Water Resour. Res.*, *35*(5), 1395–1407.
Ginn, T. R., and F. J. Loge (2007), Dose-structured population dynamics, *Math. Biosci.*, *208*(1), 325–343.
Goode, D. J. (1996), Direct simulation of groundwater age, *Water Resour. Res.*, *32*(2), 289–296.
Gurtin, M. E., and R. C. MacCamy (1974), Non-linear age-dependent population dynamics, *Arch. Ration. Mech. Anal.*, *54*(3), 281–300.
Gurtin, M. E., and R. C. MacCamy (1979), Some simple models for nonlinear age-dependent population dynamics, *Math. Biosci.*, *43*(3), 199–211.
Harman, C. J. (2014), Time-variable transit time distributions and transport: Theory and application to storage-dependent transport of chloride in a watershed, *Water Resour. Res.*, *51*, 1–30, doi:10.1002/2014WR015707.
Isham, V., D. Cox, I. Rodriguez-Iturbe, A. Porporato, and S. Manfreda (2005), Representation of space–time variability of soil moisture, *Proc. R. Soc. A*, *461*(2064), 4035–4055.
Keyfitz, B. L., and N. Keyfitz (1997), The mckendrick partial differential equation and its uses in epidemiology and population study, *Math. Comput. Modell.*, *26*(6), 1–9.
Kirchner, J. W., X. Feng, and C. Neal (2000), Fractal stream chemistry and its implications for contaminant transport in catchments, *Nature*, *403*(6769), 524–527.
Kot, M. (2001), *Elements of Mathematical Ecology*, Cambridge Univ. Press, Cambridge, U. K.
Kottegoda, N., and R. Rosso (1997), *Statistics, Probability, and Reliability for Civil and Environmental Engineers, Civ. Eng. Ser.*, McGraw-Hill, N. Y.
Manzoni, S., G. G. Katul, and A. Porporato (2009), Analysis of soil carbon transit times and age distributions using network theories, *J. Geophys. Res.*, *114*, G04025, doi:10.1029/2009JG001070.
McDonnell, J. J., and K. Beven (2014), Debates—The future of hydrological sciences: A (common) path forward? a call to action aimed at understanding velocities, celerities and residence time distributions of the headwater hydrograph, *Water Resour. Res.*, *50*, 5342–5350, doi:10.1002/2013WR015141.
McGuire, K. J., and J. J. McDonnell (2006), A review and evaluation of catchment transit time modeling, *J. Hydrol.*, *330*(3), 543–563.
M'Kendrick, A. (1925), Applications of mathematics to medical problems, *Proc. Edinburgh Math. Soc.*, *44*, 98–130.
Murray, J. (2002), Mathematical biology, in *Interdisciplinary Applied Mathematics*, edited by S. S. Antman, et al. vol. 1, Springer, N. Y.
Nauman, E. (1969), Residence time distribution theory for unsteady stirred tank reactors, *Chem. Eng. Sci.*, *24*(9), 1461–1470.
Nauman, E. B. (2008), Residence time theory, *Ind. Eng. Chem. Res.*, *47*(10), 3752–3766.
Niemi, A. J. (1977), Residence time distributions of variable flow processes, *Int. J. Appl. Radiat. Isot.*, *28*(10), 855–860.
Pareto, V. (1964), *Cours d'Economie Politique*, Librairie Droz., Genève.
Pope, S. (2000), *Turbulent Flows*, Cambridge Univ. Press, Cambridge, U. K.
Porporato, A., G. Vico, and P. A. Fay (2006), Superstatistics of hydro-climatic fluctuations and interannual ecosystem productivity, *Geophys. Res. Lett.*, *33*, L15402, doi:10.1029/2006GL026412.
Rodhe, A., L. Nyberg, and K. Bishop (1996), Transit times for water in a small till catchment from a step shift in the oxygen 18 content of the water input, *Water Resour. Res.*, *32*(12), 3497–3511.
Rotenberg, M. (1972), Theory of population transport, *J. Theor. Biol.*, *37*(2), 291–305.
Singh, V. (1992), *Elementary Hydrology*, Prentice Hall, N. J.
Sornette, D. (2006), *Critical Phenomena in Natural Sciences: Chaos, Fractals, Selforganization and Disorder: Concepts and Tools, Springer Ser. Synergetics*, Springer, Berlin.
Stratonovich, R. L. (1967), Topics in the theory of random noise (Gordon and Breach, N. Y. 1967). *Vol. II*, 234.
Suweis, S., A. Rinaldo, S. Van der Zee, E. Daly, A. Maritan, and A. Porporato (2010), Stochastic modeling of soil salinity, *Geophys. Res. Lett.*, *37*, L07404, doi:10.1029/2010GL042495.
Suweis, S., A. Porporato, A. Rinaldo, and A. Maritan (2011), Prescription-induced jump distributions in multiplicative Poisson processes, *Phys. Rev. E*, *83*(6), 061119.
Trucco, E. (1965), Mathematical models for cellular systems the von Foerster equation. Part I, *The Bull. Math. Biophys.*, *27*(3), 285–304.
Tsurui, A., and S. Osaki (1976), On a first-passage problem for a cumulative process with exponential decay, *Stochastic Processes Appl.*, *4*(1), 79–88.
van der Velde, Y., P. Torfs, S. Zee, and R. Uijlenhoet (2012), Quantifying catchment-scale mixing and its effect on time-varying travel time distributions, *Water Resour. Res.*, *48*, W06536, doi:10.1029/2011WR011310.
Varni, M., and J. Carrera (1998), Simulation of groundwater age distributions, *Water Resour. Res.*, *34*(12), 3271–3281.
Viola, F., E. Daly, G. Vico, M. Cannarozzo, and A. Porporato (2008), Transient soil-moisture dynamics and climate change in Mediterranean ecosystems, *Water Resour. Res.*, *44*, W11412, doi:10.1029/2007WR006371.
von Foerster, H. (1959), Some remarks on changing populations, in *The Kinetics of Cellular Proliferation*, edited by J. F. Stohlman, pp. 382–407, Grune and Stratton, N. Y.
Weiss, G. (1977), Shot noise models for the generation of synthetic streamflow data, *Water Resour. Res.*, *13*(1), 101–108.