A storage model with random release rate for modeling exposure to food contaminants

Patrice Bertail
CREST (INSEE) - Laboratoire de Statistiques

Stéphan Clémençon
MODALX - Université Paris X Nanterre
&
Unité Métarisk - Institut National de la Recherche Agronomique

Jessica Tressou
Department of Information and Systems Management
Hong Kong University of Science and Technology

September 5, 2006

Abstract

This paper is devoted to present and study a specific continuous-time piecewise-deterministic Markov process for describing the temporal evolution of exposure to a given food contaminant. The quantity X of food contaminant present in the body evolves through its accumulation after repeated dietary intakes on the one hand and the pharmacokinetics behavior of the chemical on the other hand. In the dynamic modeling considered here, the accumulation phenomenon is modeled by a simple marked point process with positive i.i.d. marks and elimination in between intakes occurs at a random linear rate θX, randomness of the coefficient θ accounting for the variability of the elimination process due to metabolic factors. Via embedded chain analysis, ergodic properties of this extension of the standard compound Poisson dam with (deterministic) linear release rate are investigated, the latter being of crucial importance for describing the long-term behavior of the exposure process $(X_t)_{t\geq 0}$ and assessing values of quantities such as the proportion of time the body burden in contaminant is over a certain threshold. The exposure process being not directly observable, simulation-based statistical methods for estimating steady-state or time-dependent quantities are also investigated by coupling analysis. Finally, applications to methylmercury contamination data are considered.

1 Introduction

Certain foods may contain varying amounts of chemicals such as methylmercury (present in sea food), dioxins (in poultry, meat) or mycotoxins (in cereals, dried fruits, etc.), which may cause major health problems when accumulating inside the body in excessive doses. Food safety is now a crucial stake as regards public health in many countries (as an example, it is a thematic top priority of the 7th European Research Framework program, see http://ec.europa.eu/research/fp7/). This topic naturally interfaces with various disciplines, such as biology, nutritional medicine, toxicology and of course applied mathematics with the aim to develop rigorous methods for quantitative risk assessment. A scientific literature devoted to probabilistic and statistical methods for the study of dietary exposure to food contaminants is progressively carving out a place in applied probability and statistics journals (see [42], [18], [23] or [7] for instance).
Static viewpoints for the probabilistic modeling of the quantity X of a given food contaminant ingested on a short period have been considered in recent works, mainly focussing on the tail behavior of X and allowing for computation of the probability that X exceeds a maximum tolerable dose (see [6], [41]). However, such approaches for food risk analysis do not take into account the accumulating and eliminating processes occurring in the body, which naturally requires to introduce time as a crucial description parameter of a comprehensive model.

This paper aims at proposing a dynamic modeling of exposure to a certain food contaminant, incorporating important features of the phenomenon, in particular in a way that the model may account for the contaminant pharmacokinetics in man following intakes. The case of methylmercury food contamination shall serve as a running illustration of the concepts and methods studied in this article: mathematical modeling of the pharmacokinetics behavior in man of methylmercury (essentially present in sea food products) has received increasing attention in the toxicology literature (see [31], [39] [40], [1] or [20]) and dose-response relationships have been extensively investigated for this contaminant, establishing clearly its negative impact on human health (refer to [13]). In our modeling the amount of contaminant present in the body evolves through its accumulation after repeated intakes (food consumption) and according to the pharmacokinetics governing its elimination/excretion, so that its temporal evolution is described by a piecewise-deterministic Markov process (*PDM process* in abbreviated form): the accumulation process is modeled by a marked point process in a standard fashion, while the elimination phenomenon is described by a differential equation with random coefficients, randomness accounting for the variability of the rate at which the total contaminant body burden decreases in between intakes due to metabolic factors. Such a process slightly extends storage models with general release rules widely used in operations research and

engineering for dealing with problems such as water storage in dams, in that one allows here the (content dependent) release rate to be random, as strongly advocated by biological modeling, and inter-intake times are not required to be exponentially distributed.

The outline of the paper is as follows. In section 2 a class of stochastic models with a reasonably simple (markovian) structure for describing the evolution through time of food contaminant exposure is introduced. In the important case when the (random) elimination rate is linear (such a feature being strongly motivated by previous works on kinetics modeling), ergodic properties of the *exposure process* are thoroughly investigated in section 3. As the exposure process cannot be observed in general, practical statistical inference techniques based on simulation methods for estimating steady-state or time-dependent quantities in this model are presented and studied in section 4. Finally, empirical studies related to methylmercury food contamination are carried out in section 5, with the aim to illustrate the relevance of the modeling and the statistical methods studied in this paper.

2 Modeling the exposure to a food contaminant

Suppose that an exhaustive list of P types of food, indexed by $p = 1, \ldots, P$, involved in the alimentation of a given population and possibly contaminated by a certain chemical, is drawn up. Each type of food $p \in \{1, \ldots, P\}$ is contaminated in random ratio $K^{(p)}$, with probability distribution $F_{\mathcal{K}^{(p)}}$, regarding to the chemical of interest. Concerning this specific contaminant exposure, a meal may be viewed as a realization of a r.v. $Q = (Q^{(1)}, \ldots, Q^{(P)})$ indicating the quantity of food of each type consumed, renormalized by the body weight. For a meal Q drawn from a distribution $F_{\mathcal{Q}}$ on $(\mathbb{R}_+^P, \mathcal{B}_{\mathbb{R}_+^P})$, cooked from foods of which toxicity is described by a contamination ratio vector $K = (K^{(1)}, \ldots, K^{(P)})$ drawn from $F_{\mathcal{K}} = \otimes_{p=1}^{P} F_{\mathcal{K}^{(p)}}$, the global contaminant intake is

$$U = \langle K, Q \rangle , \tag{1}$$

denoting by $\langle ., . \rangle$ the standard inner product on $\mathbb{R}^P$. Its probability distribution $F_{\mathcal{U}}$ is the image of $F_{\mathcal{K}} \otimes F_{\mathcal{Q}}$ by the inner product $\langle ., . \rangle$, assuming that the quantities of food consumed are independent from the contamination levels. Here and throughout, we suppose that the contaminant intake distribution F_U has a density f_U with respect to λ, the Lebesgue measure on $\mathbb{R}_+$.

By convention, $T_0 = 0$ is chosen as time origin. The food contamination phenomenon through time for an individual of the population of interest may be classically modeled by a marked point process $\{(T_n, Q_n, K_n)\}_{n \geq 1}$ on $\mathbb{R}_+ \times \mathbb{R}_+^P \times \mathbb{R}_+^P$, the T_n's being the successive times at which the individual consumes foods among the list $\{1, \ldots, P\}$ and the marks (Q_n, K_n) being respectively

the vector of food quantities and the vector of contamination ratios related to the meal had at time T_n, $n \geq 1$. The process $\{(T_n, Q_n)\}_{n\geq 1}$ describing dietary behavior is assumed independent from the sequence $(K_n)_{n\geq 1}$ of chemical contamination ratios. Although the modeling of dietary behaviors could certainly give rise to a huge variety of models, depending on the dependence structure between (T_n, Q_n) and past values $\{(T_m, Q_m)\}_{m<n}$ that one stipulates, we make here the simplifying assumption that the marks Q_n, $n \geq 1$, form an i.i.d. sequence with common distribution F_Q, independent from the location times $(T_n)_{n\geq 1}$. This assumption is acceptable for chemicals present in a few types of food, such as methylmercury, our running example, but certainly not for all contaminants. For chemicals present in many foods of everyday consumption such as Ochratoxin A (present in cereals, coffee, *etc.*), it would be necessary to introduce additional autoregressive structure in the model for capturing important features of any realistic diet (the consumption of certain food being typically alternated for reasons related to taste or nutritional aspects). Such a modeling task is beyond the scope of the present paper and is left for further investigation. Finally, we suppose that the inter-intake times $\Delta T_{n+1} = T_{n+1} - T_n$, $n \geq 1$, form a sequence of i.i.d. r.v.'s with common probability distribution $G(dt) = g(t)dt$ and finite expectation $m_G = \int_{t=0}^{\infty} tG(dt) < \infty$, the sequence $(T_n)_{n\geq 1}$ of intake times being thus a pure renewal process.

Contamination sources other than dietary are neglected in the present study and we denote by $X(t)$ the total body burden in contaminant at time $t \geq 0$. In between intakes, we assume that the *contamination exposure process* $X(t)$ is governed by the differential equation

$$\dot{x}(t) = -r(x(t), \theta), \tag{2}$$

θ being a random parameter, taking its values in a set $\Theta \subset \mathbb{R}^d$ with $d \geq 1$ say, and accounting in the modeling for fluctuations of the (content dependent) elimination rate due to metabolic factors at the intake times (the successive values θ_n, $n \in \mathbb{N}$, of θ are thus fixed at times $T_0, T_1, \ldots$). And the function $r(x, \theta)$ is assumed to be strictly positive and continuous on $\mathbb{R}_+^* \times \Theta$, such that for all $\theta \in \Theta$, $r(0, \theta) = 0$ (so that when $X(t)$ eventually reaches the level 0, the process stays at this level until the next intake) and for all $(\epsilon, \theta) \in (0, 1) \times \Theta$:

$$\inf_{\epsilon<x<\epsilon^{-1}} r(x, \theta) > 0 \text{ and } \sup_{0<x<\epsilon^{-1}} r(x, \theta) < \infty. \tag{3}$$

Under these conditions, for any initial value $x(0) \geq 0$ and metabolic parameter value $\theta \in \Theta$, Eq. (2) has clearly a unique solution.

Remark 1 In toxicology, Eq. (2) is widely used with $r(x, \theta) = \theta x$ for modeling the kinetics in man of certain contaminants following intakes. As shown by many pharmacokinetics studies, there is considerable empirical evidence that it properly describes the way the elimination rate depends on the total body

burden of the chemical in numerous cases (see [39] and the references therein). In this context, the release parameter $\log 2/\theta$ is known as the *half-life* of the contaminant in the body (the time required for X to decrease by half in absence of new contaminant intake).

Remark 2 Other approaches may be naturally adopted for describing the elimination phenomenon occurring in between intakes. For instance, toxicokinetic models based on stochastic differential equations or decreasing jump processes (as in inventory modeling) could be pertinently considered for this purpose.

We assume that $(\theta_n)_{n\in\mathbb{N}}$ is an i.i.d. sequence with common distribution $H(d\theta)$. For a given value of the metabolic parameter $\theta \in \Theta$, the time necessary for the body burden (without further intake) to decrease from $x_0 > 0$ to $x \in (0, x_0)$ is given by

$$\tau_\theta(x_0, x) = \int_x^{x_0} \frac{1}{r(y,\theta)}\, dy.$$

Under the assumptions stated above, we clearly have that $H(\{\tau_\theta(x_0, x) < \infty\}) = 1$ for all $0 < x \leq x_0$. The contaminant may be thus entirely eliminated from the body (the amount x reaching then the level 0) with probability one in the sole case when the following condition holds.

Condition ($\mathcal{C}_1$): $H(\{\tau_\theta(x_0, 0) < \infty\}) = 1$ for some $x_0 > 0$.

In such a case we would also have $H(\{\tau_\theta(x, 0) < \infty\}) = 1$ for all $x \geq 0$. In this respect, it is noteworthy that, in the linear case mentioned in Remark 2, $\tau_\theta(x, 0) = \infty$ for all $\theta > 0$ and $x > 0$.

Hence, in between intake times and given the current value of the metabolic parameter θ, the process moves in a deterministic fashion according to (2), and has the same (upward) jumps as the process of cumulative intakes

$$B(t) = \sum_{n=1}^{N(t)} U_n, \tag{4}$$

with $U_n = \langle K_n, Q_n \rangle$, $n \in \mathbb{N}$, and denoting by $N(t) = \sum_{n\in\mathbb{N}} \mathbb{I}\{T_n \leq t\}$ the number of intakes until time t. The process X is piecewise-deterministic with càd-làg trajectories (see a typical sample path in Fig.1) and satisfies the equation

$$X(t) = X(0) + B(t) - \sum_{n=1}^{N(t)+1} \int_{T_{n-1}}^{T_n \wedge t} r(X(s), \theta_n)\, ds, \tag{5}$$

$X(0)$ denoting the total body burden in contaminant at initial time $T_0 = 0$. For an account of such piecewise deterministic processes, one may refer to [17] (see also [16] and ergodic results may be found in [12]).

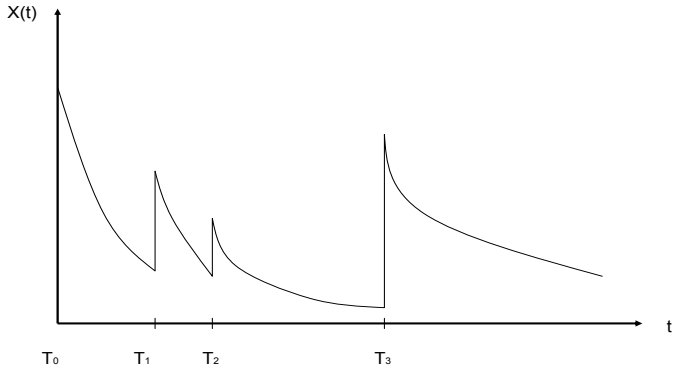

Figure 1: Sample path of the exposure process X.

For the continuous-time process thus defined to be markovian, one has to record the current value $\theta(t) = \sum_{n\in\mathbb{N}} \theta_n \mathbb{I}_{\{t\in[T_n,T_{n+1}[\}}$ of the metabolic parameter as well as the backward recurrence time $A(t) = t - T_{N(t)}$ (the time since the last intake). By construction, the process $(X(t), \theta(t), A(t))_{t\geq 0}$ is strongly Markovian with generator

$$\mathcal{G}\phi(x,\theta,t) = \zeta(t)\int_{u=0}^{\infty}\int_{\theta'\in\Theta}\{\phi(x+u,\theta',0)-\phi(x,\theta,t)\}F_U(du)H(d\theta') - r(x,\theta)\partial_x\phi(x,\theta,t) + \partial_t\phi(x,\theta,t), \quad (6)$$

denoting by $\zeta(t) = g(t)/\int_{s=t}^{\infty} g(s)ds$ the hazard rate of the inter-intake times and provided that $\phi(.,\theta,.) : (x,t) \mapsto \phi(x,\theta,t)$ is a bounded function with bounded continuous first derivatives in x and t for all $\theta \in \Theta$.

In the above setting, the time origin $T_0 = 0$ does not necessarily correspond to an intake time. Given the time $A(0) = a$ since the last intake at time $t = 0$, we let ΔT_1 have the density $g_a(t) = g(a+t)/\int_{s=a}^{\infty} g(s)ds$, making the renewal process $(\Delta T_n)_{n\in\mathbb{N}}$ possibly delayed, except naturally in the case when the inter-intake distribution G is exponential. However, the choice of such a memoryless distribution in the dietary context is clearly not pertinent, distributions with increasing hazard rate being more adequate. Here and throughout we denote by $\mathbb{P}_{x,a}$ the probability measure on the underlying space such that $(X(0), A(0)) = (x, a)$ and $\theta(0) \sim H$, and by $\mathbb{E}_{x,a}(.)$ the $\mathbb{P}_{x,a}$-expectation for all $x \geq 0$ and $a \in \text{supp}(G)$.

In the case when one neglects variability in the elimination process (*i.e.* when H is a Dirac measure), this modeling boils down to a standard *storage model with a general release rate* (see [11] and [10] for instance). We refer to Chapter XIV in [3] for an account of such processes, widely used in operations research for modeling queuing/storage systems. Basic communication and stochastic stability properties of the stochastic process $X = (X(t))_{t\geq 0}$ may be established in a fashion very similar to the ones of the latter processes, although the

additional assumption that the renewal times are exponentially distributed is usually required in these studies, making the process X itself Markovian (which facilitates much the study but is not relevant to our application as emphasized above). They are summarized in the next result (of which proof is omitted since it is a slight modification of the proof of Proposition 1.2 in chap. XIV of [3]).

Theorem 1 *Suppose that* $G(dx) = g(x)dx$ *has infinite tail. Assume further that either* $g(x) > 0$ *on* $]0, \epsilon]$ *for some* $\epsilon > 0$ *or else that* $F_{\mathcal{U}}$ *has infinite tail. Then* X *reaches any state* $x > 0$ *in finite time with positive probability whatever the starting point, i.e. for all* $x_0 \geq 0$, $a \in \mathrm{supp}(G)$, *it holds*

$$\mathbb{P}_{x_0,a}(\tau_x < \infty) > 0, \tag{7}$$

with $\tau_x = \inf\{t \geq 0 : X_t = x\}$. *Furthermore, if condition* $(\mathcal{C}_1)$ *is fulfilled, then (7) still holds for* $x = 0$.
Besides, either X *"heads to infinity" with probability one, i.e. is such that* $\mathbb{P}_{x_0,a}(\{X(t) \to \infty \text{ , as } t \to \infty\}) = 1$ *for all* $x_0 \geq 0$, *or else* X *reaches any state* $x > 0$ *in finite time with probability one whatever the starting point, i.e. for all* $x_0 \geq 0$, $a \in \mathrm{supp}(G)$,

$$\mathbb{P}_{x_0,a}(\tau_x < \infty\}) = 1. \tag{8}$$

If $(\mathcal{C}_1)$ *is satisfied, then (8) also holds for* $x = 0$.

An important task is to find conditions ensuring that the limiting behavior of the exposure process X is represented by a stationary probability measure μ describing the equilibrium state to which the process settles as time goes to infinity. In particular, time averages over long periods, such as the mean time spent by the exposure process X over a threshold $u > 0$, $T^{-1}\int_0^T \mathbb{I}_{\{X_t \geq u\}}dt$, for instance, are then asymptotically described by the distribution μ. Beyond stochastic stability properties, evaluating the rate at which the process converges to the stationary state is also of critical importance in practice. These questions are thoroughly investigated for linear rate models in the next section.

3 Probabilistic study in the linear rate case

We now focus on ergodicity properties of the exposure process $X(t)$ in the specific case when for a given metabolic state described by a real parameter θ, the elimination rate is proportional to the total body burden in contaminant, *i.e.*

$$r(x, \theta) = \theta x. \tag{9}$$

Here we suppose that Θ is a subset of $\mathbb{R}_+^*$, ensuring that (3) is satisfied. As mentioned before, the linear case is of crucial importance in toxicology, insofar as it suitably models the pharmacokinetics behavior in man of numerous

chemicals. We shall show that studying the long-term behavior of X boils down to investigating the properties of the embedded Markov chain $\tilde{X} = (X_n)_{n\geq 1}$ of which values correspond to the ones taken by the exposure process just after intake times : $X_n = X(T_n)$ for all $n \geq 1$. By construction, the chain $\tilde{X}$ satisfies the following autoregressive equation with random coefficients

$$X_{n+1} = e^{-\theta_n \Delta T_{n+1}} X_n + U_{n+1}, \text{ for all } n \geq 1, \tag{10}$$

and has transition probability $\Pi(x, dy) = \pi(x, y)dy$ with transition density

$$\pi(x, y) = \int_{\theta\in\Theta} \int_{t=\frac{1}{\theta}\log(1\vee\frac{x}{y})}^{\infty} f_U(y - xe^{-\theta t})G(dt)H(d\theta), \tag{11}$$

for all $(x, y) \in \mathbb{R}_+^{*2}$, where $a \vee b = \max(a, b)$. Ergodicity of such real-valued Markov chains Y, defined through stochastic recurrence equations of the form $Y_{n+1} = \alpha_n Y_n + \beta_n$, where $\{(\alpha_n, \beta_n)\}_{n\in\mathbb{N}}$ is a sequence of i.i.d. pairs of positive r.v.'s, has been extensively studied in the literature, such models being widely used in financial or insurance mathematics (see section 8.4 in [19] for instance). Specialized to our setting, well known results related to such processes enable to show that the embedded chains $\tilde{X}$ is positive recurrent under the assumption that $\log(1 \vee U_1\})$ has finite expectation, as stated in the next theorem. Furthermore, the simple autoregressive form of Eq. (10) makes Foster-Lyapunov conditions easily verifiable for such Markov chains, in order to refine their stability analysis.

Theorem 2 *Under the assumptions of Theorem 1, the chain $\tilde{X}$ is λ- irreducible. Moreover, suppose that the following condition holds.*

- *(H1)* $\mathbb{E}[\log(1 \vee U_1)] < \infty$.

Then $\tilde{X}$ is recurrent positive with stationary probability distribution $\tilde{\mu}$. If one assume further that:

- *(H2) there exists some $\gamma \geq 1$ such that $\mathbb{E}(U_1^\gamma) < \infty$,*

then $\tilde{X}$ is geometrically ergodic, $\tilde{\mu}$ has finite expectation and there exist constants $R < \infty$ and $r > 1$ such that, for all $n \geq 1$, $x > 0$,

$$\sup_{\{\psi, |\psi(z)|\leq 1+z^\gamma\}} \left| \int_{y=0}^{\infty} \psi(y)\Pi^n(x, dy) - \tilde{\mu}(\psi) \right| \leq R(1 + x^\gamma)r^{-n}, \tag{12}$$

denoting by Π^n the n-th iterate of Π and with $\tilde{\mu}(\psi) = \int_{y=0}^{\infty} \psi(y)\tilde{\mu}(dy)$ for any $\tilde{\mu}$-integrable function ψ. Suppose finally that the next condition holds,

- *(H3) The r.v. U_1 is regularly varying with index $\kappa > 0$.*

Then the stationary law $\tilde{\mu}$ has regularly varying tail with index κ.

Remark 3

- The relevance of the regular variation assumption for modeling the tail behavior of dietary contaminant intakes related to certain chemicals is strongly supported in [41] and [6]. In these works, various estimation strategies for tail distribution features are also proposed and implemented on several food contamination and consumption data sets.
- It is noteworthy that estimates of the constants r, R involved in the rate bound (12) may be explicitly computed from the parameters of the Foster-Lyapunov drift condition fulfilled by the chain (see Theorem 5.1 in [37]).

PROOF. From conditions required by Theorem 1, aperiodicity and irreducibility properties are immediately established for the discrete-time chain $\tilde{X}$. Besides, under mild irreducibility conditions, the stability of the *random coefficients autoregressive model* on $\mathbb{R}^d$

$$Y_{n+1} = \alpha_n Y_n + \beta_n,$$

where (α_n, β_n), $n = 1, \ldots$ are i.i.d. r.v.'s on $\mathbb{R}_+^* \times \mathbb{R}^d$, has been investigated in detail since the seminal contribution of [27] (see [34] and the references therein). Under the assumption that $\mathbb{E}[\log(1 \vee \|\beta_1\|)] < \infty$ and $\mathbb{E}[\log(1 \vee \alpha_1)] < \infty$, it is well known that a sufficient and necessary condition for the chain X to have a (unique) probability measure is that $\mathbb{E}[\log(\alpha_1)] < 0$ (see Corollary 2.7 in [9] for instance). Based on this result, it is then straightforward that, under (H1) and (H2), the chain $\tilde{X}$ is positive recurrent with absolutely continuous stationary probability distribution $\tilde{\mu}(dx) = \tilde{f}(x)dx$.
In the discrete-time context, analysis of the stability of Markov models $(Y_n)_{n\in\mathbb{N}}$ may be carried out by establishing suitable conditions for the 'drift' $\Delta V(y) = \mathbb{E}[V(Y_1) \mid Y_0 = y] - V(y)$ for appropriate non-negative test functions $V(y)$. Such 'Foster-Lyapunov' criteria stipulate the existence of a 'small set' S (*i.e.* an accessible set S to which the chain returns in a given number of steps with positive probability, uniformly bounded by below, see section 5.2 in [30]) towards which the chain drifts in the sense that:

$$\Delta V(x) \leq -f(x) + b\mathbb{I}_{\{x\in S\}}, \tag{13}$$

for some 'norm-like' function $f(x) \geq 1$ and $b < \infty$. Now for the chain $\tilde{X}$, any compact interval $[0, x_0]$, with $x_0 > 0$ large enough, is small (one may find $\delta(x_0) > 0$ such that: $\forall x \in [0, x_0]$, $\Pi(x, [0, x_0]) \geq \delta(x_0)$). When $\gamma = 1$ for instance, take $V(x) = 1 + x$. The affine drift related to $\tilde{X}$ is given by

$$\Delta V(x) = -cx + \mathbb{E}(U_1),$$

with $c = 1 - \mathbb{E}(e^{-\theta_1 \Delta T_2}) > 0$. Choosing $S = [0, s]$ with $s \geq 2/c\mathbb{E}(U_1) + c$, (13) is fulfilled with $f(x) = cV(x)/2$ and $b = \mathbb{E}(U_1) + c/2$. Applying Theorem 15.0.1

in [30], we thus get that $\tilde{X}$ is geometrically ergodic with invariant probability measure $\tilde{\mu}$ such that $\tilde{\mu}(V) = \int_{x=0}^{\infty} V(x)\tilde{\mu}(dx) < \infty$. In particular, $\tilde{\mu}$ has finite expectation and there exist constants $r > 1$, $R < \infty$ such that for all $x > 0$:

$$\sum_{n=0}^{\infty} r^n \left\| \Pi^n(x,.) - \tilde{\mu} \right\|_V \leq RV(x), \tag{14}$$

with $\|\nu\|_V = \sup_{\psi:|\psi|\leq V} \left| \int \psi(x)\nu(dx) \right|$ for all bounded measure ν on the real line. When $V \equiv 1$, $\|.\|_V$ coincides with the total variation norm $\|.\|_{TV}$. For $\gamma > 1$, the results is proved in a similar fashion by taking $V(x) = 1 + x^{\gamma}$.
Finally, the last assertion of Theorem 2 immediately derives from Theorem 1 in [25]. ■

As pointed out in [29], stochastic stability analysis based on drift criteria in the continuous-time setting is not as straightforward as in the discrete-time case, generally due to the complex form of the generator and of candidate test functions. However, given the explicit relationship between X and the embedded discrete-time chain $\tilde{X}$ in our specific case, the properties of the continuous-time model may be investigated based on the results established above for $\tilde{X}$ and on further moment conditions for the inter-intake distribution, as the one below.

- (H4) There exists $\delta > 0$ such that $\mathbb{E}[\exp(\delta \Delta T_2)] < \infty$.

Theorem 3 *Under the assumptions of Theorem 1 and supposing that* (H_1) *is fulfilled,* $X(t)$ *has an absolutely continuous limiting probability distribution* μ *given by*

$$\mu([u,\infty[) = m_G^{-1} \int_{x=u}^{\infty} \int_{t=0}^{\infty} \int_{\theta\in\Theta} t \wedge \frac{\log(x/u)}{\theta} \tilde{\mu}(dx)G(dt)H(d\theta), \tag{15}$$

in the sense that $T^{-1}\int_0^T \mathbb{I}_{\{X_t \leq u\}} dt \to \mu(]0,u])$, $\mathbb{P}_{x_0,a}$*-a.s., as* $t \to \infty$ *for all* $x_0 \geq 0$ *and* $a \in \mathrm{supp}(G)$.

Furthermore,

- *if* (H3) *holds and the set* Θ *is bounded, then* μ *is regularly varying with the same index as* F_U,

- *and if* (H2) *and* (H4) *hold and* G *has finite variance* σ_G^2*, then* μ *has finite moment of order* γ *and for all* $(x,a) \in \mathbb{R}_+^* \times \mathrm{supp}(G)$ *there exist constants* $k \in]0,1[$, $K_a < \infty$ *such that*

$$\sup_{\psi(z)\leq 1+z^{\gamma}} |\mathbb{E}_{x,a}[\psi(X_t)] - \mu(\psi)| \leq K_a(1+x^{\gamma})k^t. \tag{16}$$

Remark 4 When the U_n's are heavy-tailed, and under the assumption that the ΔT_n's are exponentially distributed (making $B(t)$ a time-homogeneous Lévy

process), the fact that the stationary distribution μ inherits its tail behavior from F_U has been established in [2] for general deterministic release rates. Besides, when assuming G exponential and θ fixed, one may identify the limit distribution μ in some specific cases (see section 8 in [10] or section 2 in Chap. XIV of [3]) using basic level crossing arguments (X being itself markovian in this case). If F_U is also exponential for instance, μ is a Gamma distribution. And furthermore, due to the simple form of the generator in the latter case, one may establish an exponential rate of convergence to μ by standard drift criterion or coupling arguments (see section 5 in [38]).

PROOF. Set $X_0 = X(0)$. Observe that for all $t > 0$,

$$X(t) = X_{N(t)} e^{-\theta_{N(t)} A(t)}, \tag{17}$$

so that $X(t) \leq X_{N(t)}$. Hence we naturally have $\{X(t) \to \infty\} \subset \{X_n \to \infty\}$. Therefore, under (H_1), we know that $\tilde{X}$ is positive recurrent with stationary distribution $\tilde{\mu}$, so that in particular $\mathbb{P}(X_n \to \infty) = 0$. Furthermore, observe that for all $t > 0$, $u \geq 0$:

$$\int_{s=0}^{t} \mathbb{I}_{\{X(s)\geq u\}} ds = \sum_{k=1}^{N(t)} \int_{s=T_{k-1}}^{T_k} \mathbb{I}_{\{X(s)\geq u\}} ds + \int_{s=T_{N(t)}}^{t} \mathbb{I}_{\{X(s)\geq u\}} ds. \tag{18}$$

Therefore, for all $k \in \mathbb{N}$

$$\int_{s=T_k}^{T_{k+1}} \mathbb{I}_{\{X(s)\geq u\}} ds = \mathbb{I}_{\{X_k \geq u\}} \cdot \Delta T_{k+1} \wedge \frac{\log(X_k/u)}{\theta_k}. \tag{19}$$

Now, applying the SLLN to the positive recurrent chain $((X_n, \theta_n, \Delta T_{n+1}))_{n\in\mathbb{N}}$ with invariant probability distribution $\tilde{\mu}(dx) \otimes H(d\theta) \otimes G(dt)$, we get that

$$n^{-1} \sum_{k=1}^{n} \int_{s=T_k}^{T_{k+1}} \mathbb{I}_{\{X(s)\geq u\}} ds \to \int_{x=u}^{\infty} \int_{\theta\in\Theta} \int_{t=0}^{\infty} t \wedge \frac{\log(x/u)}{\theta} \tilde{\mu}(dx) H(d\theta) G(dt).$$

As we assumed $m_G = \mathbb{E}(\Delta T_k) < \infty$ for $k \geq 2$, we have the following convergence for the delayed renewal process: $N(t)/t \to m_G^{-1}$ as $t \to \infty$. Combined with (3), this yields $t^{-1} \int_{s=0}^{t} \mathbb{I}_{\{X(s)\geq u\}} ds \to \mu([u, \infty[)$ as $t \to \infty$, with μ given by (15). We thus proved that $X(t)$ has a limiting probability distribution μ, which has density $f(y)$ given by

$$f(y) = m_G^{-1} \int_{\theta\in\Theta} \int_{t=0}^{\infty} \tilde{f}(ye^{\theta t}) e^{\theta t} \bar{G}(t) dt H(d\theta),$$

denoting by $\bar{G} = 1 - G$ the inter-intake survival function.
Besides, if $\sup\Theta < \infty$, from (15) we immediately have that, for all $u > 0$, $t > 0$,

$$\frac{t \wedge \log 2}{m_G \sup\Theta} \bar{G}(t) \tilde{\mu}([2u, \infty[) \leq \mu([u, \infty[) \leq \tilde{\mu}([u, \infty[).$$

The distributions μ and $\tilde{\mu}$ have thus exactly the same right tail behavior.
We now turn to establish a rate bound for the convergence of the $X(t)$'s distribution to μ. For simplicity's sake, suppose that the renewal process $(\Delta T_n)_{n\in\mathbb{N}}$ is zero-delayed, *i.e.* T_0 is an intake time. Extension of the argument below to the delayed case when $\Delta T_1 \sim g_a(t)dt$ for some $a > 0$ is straightforward. The backward recurrence time $A(t)$ has then the distribution $Q_t(ds)$ supported by $(0,t)$ with density

$$q_t(s) = \sum_{k\in\mathbb{N}} g^{*k}(t-s)\bar{G}(s),$$

denoting by g^{*k} the k-fold convolution power of the density g. As G has finite mean m_G by assumption, $A(t)$'s distribution converges to the equilibrium distribution $Q_\infty(ds)$ with density $q_\infty(s) = m_G^{-1}\bar{G}(s)$ (refer to Chap. XI in [22] for basics in renewal theory). Furthermore, under (H4) we have by virtue of Theorem 4.4 in [28] that there exist constants $D < \infty$, $\rho \in]0,1[$ s.t.

$$\|Q_t - Q_\infty\|_{TV} = \int_{s=0}^{\infty} |q_t(s) - q_\infty(s)|\, ds \leq D\rho^t, \tag{20}$$

for all $t \geq 0$. Besides, observe that, for all $y > 0$, we have

$$\begin{aligned}\mathbb{P}_{x,0}(X(t) < y) &= \sum_{k\in\mathbb{N}} \mathbb{P}_x(X_k e^{-\theta_k(t-T_k)} \geq y, T_k \leq t, T_{k+1} > t)\\ &= \sum_{k\in\mathbb{N}} \int_{\theta\in\Theta}\int_{s=0}^{t}\int_{\{ze^{-\theta s}<y\}} \Pi^k(x,dz)\bar{G}(s)g^{*k}(t-s)dsH(d\theta).\end{aligned}$$

The distribution $P_t(x,dy)$ of X_t when $(X(0),\theta(0),A(0)) \sim \delta_x \otimes H \otimes \delta_0$ has thus the density $p_t(x,y) = \sum_{k\in\mathbb{N}} \int_{\theta\in\Theta}\int_{s=0}^{t} \pi^k(x,e^{\theta s}y)e^{\theta s}\bar{G}(s)g^{*k}(t-s)dsH(d\theta)$. For all $t > 0$, define $\bar{f}_t(y) = \int_{\theta\in\Theta}\int_{s=0}^{t} \tilde{f}(ye^{\theta s})e^{\theta s}q_t(s)dsH(d\theta)$ and observe that, for all $K \in \mathbb{N}$, one may write

$$\begin{aligned}p_t(x,y) - f(y) = &\sum_{k\leq K} \int_{\theta\in\Theta}\int_{s=0}^{t} \{\pi^k(x,e^{\theta s}y) - \tilde{f}(e^{\theta s}y)\}e^{\theta s}\bar{G}(s)g^{*k}(t-s)dsH(d\theta)\\ &+ \sum_{k>K} \int_{\theta\in\Theta}\int_{s=0}^{t} \{\pi^k(x,e^{\theta s}y) - \tilde{f}(e^{\theta s}y)\}e^{\theta s}\bar{G}(s)g^{*k}(t-s)dsH(d\theta)\\ &+ \bar{f}_t(y) - f(y).\end{aligned}$$

Let $\psi : \mathbb{R}_+^* \to \mathbb{R}$ be a Borelian function s.t. $\psi(y) \leq V(y) = 1 + y^\gamma$ for all $y > 0$. From (14) and the decomposition above combined with straightforward changes of variables, we deduce that

$$\left|\int_{s=0}^{\infty} \psi(y)\{p_t(x,y) - f(y)\}dy\right| \leq I_1(t) + I_2(t) + I_3(t), \tag{21}$$

with

$$I_1(t) = RV(x)\mathbb{P}(T_{K+1} > t),$$

$$I_2(t) = RV(x)r^{-K},$$
$$I_3(t) = \int_{y=0}^{\infty} V(y)\tilde{\mu}(dy)\,\|Q_t - Q_\infty\|_{TV}\,.$$

Therefore $\mathbb{P}(T_{K+1} > t) = \mathbb{P}(\delta \sum_{j=1}^{K} \Delta T_j > \delta t) \leq e^{-\delta t + K\log(C_\delta)}$ where $C_\delta = \mathrm{E}\exp(\delta \Delta T_j) < \infty$. Choosing $K \sim \frac{\delta t}{2\log(C_\delta)}$ yields $\mathbb{P}(T_{K+1} > t) \vee r^{-K} \leq Ck^t$ for some well chosen constants $C < \infty$ and $k \in]0,1[$. Now combined with (21) and (20), this establishes (16). ■

In order to exhibit connections between the exposure process $X = (X(t))_{t\geq 0}$ and possible negative effects of the chemical on human health, it is appropriate to consider simple characteristics of the process X, easily interpretable from an epidemiology viewpoint. In this respect, the mean exposition over a long time period $T^{-1}\int_{t=0}^{T} X(t)dt$ is one of the most relevant features. Its asymptotic behavior is refined in the next result.

Proposition 4 *Under the assumptions of Theorem 1 and supposing that* (H2) *is fulfilled for* $\gamma = 1$*, we have for all* $(x_0, a) \in \mathbb{R}_+ \times \mathrm{supp}(G)$

$$\bar{X}_T = \frac{1}{T}\int_{t=0}^{T} X(t)dt \to m_\mu, \mathbb{P}_{x_0,a}\text{-}a.s.\ , \tag{22}$$

as $T \to \infty$ *with* $m_\mu = \int_{x=0}^{\infty} x\mu(dx)$*. Moreover, if* (H2) *is fulfilled with* $\gamma \geq 2$*, then there exists a constant* $0 < \sigma^2 < \infty$ *s.t. for all* $(x_0, a) \in \mathbb{R}_+ \times \mathrm{supp}(G)$ *we have the following convergence in* $\mathbb{P}_{x_0,a}$*-distribution*

$$\sqrt{T}(\bar{X}_T - m_\mu) \Rightarrow \mathcal{N}(0, \sigma^2) \ \textit{as}\ T \to \infty. \tag{23}$$

Remark 5 • As will be shown in the proof below, the asymptotic variance σ^2 in (23) may be related to the limiting behavior of a certain additive functional of the Markov chain $((X_n, \theta_n, \Delta T_{n+1}))_{n\geq 1}$. In [4] (see also [5]), an estimator of the asymptotic variance of such functionals based on pseudo-renewal properties of the underlying chain (namely, on renewal properties of a Nummelin extension of the chain) has been proposed and a detailed study of its asymptotic properties has been carried out.

• Beyond the asymptotic exposure mean or the asymptotic mean time spent by X above a certain threshold, other summary characteristics of the exposure process could be pertinently considered from an epidemiology viewpoint, among which the asymptotic tail conditional expectation $\mathbb{E}_\mu(X \mid X > u)$, denoting by $\mathbb{E}_\mu(.)$ the expectation w.r.t. μ, after the fashion of risk evaluation in mathematical finance or insurance.

PROOF. Given $(X(0), A(0)) = (x_0, a)$, we have for all $T > 0$

$$\bar{X}_T = T^{-1}\int_{t=0}^{T_1} X(t)dt + T^{-1}\sum_{k=1}^{N(T)-1}\int_{t=T_k}^{T_{k+1}} X(t)dt + T^{-1}\int_{T_{N(T)}}^{T} X(t)dt. \tag{24}$$

The first term in the right-hand side of (24) being bounded by $x_0 T_1/T$, it almost surely converges to 0 as $T \to \infty$. Besides we have for all $k \geq 1$,

$$\int_{t=T_k}^{T_{k+1}} X(t)dt = \frac{X_k}{\theta_k}(1 - e^{-\theta_k \Delta T_{k+1}})$$

Furthermore, by virtue of Theorem 2, assumption (H2) with $\gamma = 1$ ensures that $m_{\tilde{\mu}} = \int_{x=0}^{\infty} x\tilde{\mu}(dx) < \infty$ and consequently that

$$\tilde{m} = \int_{x=0}^{\infty}\int_{t=0}^{\infty}\int_{\theta\in\Theta} \frac{x(1-e^{-\theta t})}{\theta}\tilde{\mu}(dx)H(d\theta)G(dt) < \infty,$$

making the SLLN for the positive recurrent chain $((X_n, \theta_n, \Delta T_{n+1}))_{n\geq 1}$ applicable to $\sum_{n\geq 1}(1-exp(\theta_n \Delta T_{n+1}))X_n/\theta_n$ (refer to Theorem 17.3.2 in [30] for instance). We thus have that

$$N^{-1}\sum_{k=1}^{N} \frac{X_k}{\theta_k}(1 - e^{-\theta_k \Delta T_{k+1}}) \to m_{\tilde{\mu}} \int_{t=0}^{\infty}\int_{\theta\in\Theta} \frac{1-e^{-\theta t}}{\theta} H(d\theta)G(dt) \text{ a.s.,} \quad (25)$$

as $N \to \infty$. Combining (25) with $N(T)/T \to m_G^{-1}$ a.s. as $T \to \infty$, this entails that the third term in (24) tends to 0 as $T \to \infty$ and establishes (4). Notice that $m_\mu = \int_{t=0}^{\infty}\int_{\theta\in\Theta}(1 - exp(-\theta t))/\theta H(d\theta)G(dt)m_{\tilde{\mu}}/m_G$.
We now turn to the CLT's proof. Using again Theorem 2, we have that $\int x^2\tilde{\mu}(dx) < \infty$ when (H2) holds for some $\gamma \geq 2$, so that

$$\int_{x=0}^{\infty}\int_{t=0}^{\infty}\int_{\theta\in\Theta} \frac{x^2(1-e^{-\theta t})^2}{\theta^2}\tilde{\mu}(dx)H(d\theta)G(dt) < \infty$$

By virtue of the CLT for positive recurrent chains (see Theorem 17.0.1 in [30]), we have that $N^{-1/2}\sum_{k=1}^{N}\{(1-e^{-\theta_k \Delta T_{k+1}})X_k/\theta_k - \tilde{m}\}$ converges in distribution to $\mathcal{N}(0, \tilde{\sigma}^2)$ as $N \to \infty$, with

$$\tilde{\sigma}^2 = \mathbb{E}_\mu\left[\left(\frac{X_1(1-e^{-\theta_1 \Delta T_2})}{\theta_1} - \tilde{m}\right)^2\right] + 2\sum_{k=2}^{\infty}\mathbb{E}_\mu\left[\left(\frac{X_1(1-e^{-\theta_1 \Delta T_2})}{\theta_1} - \tilde{m}\right)\left(\frac{X_k(1-e^{-\theta_k \Delta T_{k+1}})}{\theta_k} - \tilde{m}\right)\right].$$

One may then easily deduce (23) from (24) with $\sigma^2 = \tilde{\sigma}^2/m_G$. ■

4 Simulation-based statistical inference

We now consider the statistical issues one faces when attempting to estimate certain features of linear rate exposure models. The main difficulty lies in the fact

that the exposure process X is generally unobservable. Food consumption data (quantities of consumed food and consumption times) related to a single individual over long time periods are scarcely available in practice. And performing measurements at all consumption times so as to record the food contamination levels appears as not easily realizable. Instead, practitioners have at their disposal some massive databases, in which information related to the dietary habits of large population samples over short periods of time is gathered. Besides, some contamination data concerning certain chemicals and types of food are stored in data warehouses and available for statistical purposes. Finally, experiments for assessing models accounting for the pharmacokinetics behavior in man of various chemicals have been carried out. And data permitting to fit values or probability distributions on the parameters of these models are consequently available. Estimation of steady-state or time-dependent features of the law $\mathcal{L}_X$ of the process X given the starting point $(X(0), A(0)) = (x_0, a) \in \mathbb{R}_+ \times \mathrm{supp}(G)$ could thus be based on preliminary computation of consistent estimates $\hat{G}$, $\hat{F}_U$ and $\hat{H}$ of the unknown df's G, F_U and H. Hence, when the quantity of interest $\mathcal{Q}(X)$ is not analytically available from (G, F_U, H), ruling out the possibility of computing plug-in estimates, a feasible method could consist in simulating sample paths starting from (x_0, a) of the approximate process $\hat{X}$ with law $\mathcal{L}_{\hat{X}}$ corresponding to the estimated df's $(\hat{G}, \hat{F}_U, \hat{H})$ and construct estimators based on the trajectories thus obtained. This leads up to investigate the *stability* of the stochastic model described in section 2 w.r.t. G, F_U and H, and consider the *continuity problem* consisting in evaluating a measure of closeness between $\mathcal{L}_X$ and $\mathcal{L}_{\hat{X}}$ making the mapping $\mathcal{L}_X \mapsto \mathcal{Q}(X)$ continuous for the functional of interest $\mathcal{Q}$ (refer to [32] for an account on this topic). Hence, convergence preservation results may be obtained via the *continuous-mapping approach* as described in [45], where it is applied to establish stochastic-process limits for queuing systems. For simplicity's sake, we take $a = 0$ in the following study and do not consider the stability issue related to the approximation of the starting point $(X(0), A(0))$, straightforward modifications of the argument below permitting to deal with the latter problem. For notational convenience, we omit to index by $(x_0, 0)$ the probabilities and expectations considered in the sequel.

Let $0 < T < \infty$. Since the exposure process X has càd-làg sample paths, we use the $\mathcal{M}_2$ topology on the Skorohod's space $D([0,T], \mathbb{R})$ induced by the Hausdorff distance on the space of completed graphs (the completed graph of $x \in D([0,T], \mathbb{R})$ being obtained by connecting $(t, x(t))$ to $(t, x(t-))$ with a line segment for all discontinuity points), allowing trajectories to be eventually close even if their jumps do not exactly match (the $\mathcal{J}_2$ topology would be actually sufficient for our purpose, refer to [26] or [45] for an account on topological concepts for sets of stochastic processes). In order to evaluate how close the approximating and true laws are, we shall establish an upper bound for the L_1-Wasserstein Kantorovich distance between the distributions $\mathcal{L}_{X(T)}$ and $\mathcal{L}_{\hat{X}(T)}$

of $X^{(T)} = (X(t))_{t\in[0,T]}$ and $\hat{X}^{(T)} = (\hat{X}(t))_{t\in[0,T]}$, which metric on the space of probability laws on $D([0,T],\mathbb{R})$ is defined as follows (refer to [33], [8]):

$$W_1^{(T)}(\mathcal{L},\mathcal{L}') = \inf_{\substack{Z'\sim\mathcal{L}'\\ Z\sim\mathcal{L}}} \mathbb{E}[m_{\mathcal{M}_2}^{(T)}(Z',Z)], \tag{26}$$

where the infimum is taken over all pairs (Z',Z) with marginals $\mathcal{L}'$ and $\mathcal{L}$ and $m_{\mathcal{M}_2}^{(T)}(Z',Z) = m_{\mathcal{H}}^{(T)}(\Gamma_{Z'},\Gamma_Z)$, denoting by $\Gamma_{Z'}$ and Γ_Z the completed graphs of Z' and Z and by $m_{\mathcal{H}}^{(T)}$ the Hausdorff metric on the set of all compact subsets of $[0,T]\times\mathbb{R}$ related to the distance $m((t_1,x_1),(t_2,x_2)) = |t_1-t_2|+|x_1-x_2|$ on $[0,T]\times\mathbb{R}$. It is well-known that this metric implies weak convergence (see [8]). As claimed in the next theorem, the law $\mathcal{L}_{\hat{X}^{(T)}}$ gets closer and closer to $\mathcal{L}_{X^{(T)}}$ as the df's $\hat{G}$, $\hat{F}_U$ and $\hat{H}$ respectively tend to G, F_U and H in the Mallows sense. For $p\in[1,\infty)$, we denote by $M_p(F_1,F_2) = (\int_0^1 \left|F_1^{-1}(t)-F_2^{-1}(t)\right|^p dt)^{1/p}$ the L_p-Mallows distance between two df's F_1 and F_2 on the real line.

Theorem 5 *Let (G,F_U,H) (resp., $(\hat{G}^{(n)},\hat{F}_U^{(n)},\hat{H}^{(n)})$ for $n\in\mathbb{N}$) be a triplet of df's on $\mathbb{R}_+$ defining a linear exposure process X (resp., $\hat{X}_{(n)}$) starting from $x_0\geq 0$ and fulfilling Theorem 1's assumptions and (H_2) with $\gamma=1$. Suppose that $M_1(\hat{G}^{(n)},G)\vee M_1(\hat{F}_U^{(n)},F_U)\vee M_1(\hat{H}^{(n)},H)\to 0$ as $n\to\infty$. Assume further that G (resp., $\hat{G}^{(n)}$) has finite variance σ_G^2 (resp., $\sigma^2_{\hat{G}^{(n)}}$) and H (resp., $\hat{H}^{(n)}$) has finite mean. If $\sigma^2_{\hat{G}^{(n)}}$ remains bounded, then:*

$$\sup_{T>0} T^{-2} W_1^{(T)}(\mathcal{L}_{X^{(T)}},\mathcal{L}_{\hat{X}^{(T)}})\to 0, \text{ as } n\to\infty. \tag{27}$$

And for all $T>0$ we have the weak convergence:

$$\hat{X}_{(n)}^{(T)} \Rightarrow X^{(T)} \text{ in } D([0,T],\mathbb{R}), \text{ as } n\to\infty. \tag{28}$$

Remark 6 • We point out that similar results hold for the L_p-Wasserstein distance with $p\in[1,\infty)$ under suitable moment conditions.

• It may also be convenient to consider the function space $D([0,\infty),\mathbb{R})$ in which X has its sample paths and on which the metric

$$m_{\mathcal{M}_2}^{(\infty)}(x,x') = \int_{t\geq 0} 2^{-t} m_{\mathcal{M}_2}^{(t)}((x_s)_{s\in[0,t]},(x'_s)_{s\in[0,t]})dt$$

for $(x,x')\in D([0,\infty),\mathbb{R})^2$ may be considered. It is noteworthy that (27) also immediately provides a control of the L_1-Wasserstein distance $W_1^{(\infty)}$ corresponding to that metric between $\mathcal{L}_X$ and $\mathcal{L}_{\hat{X}}$.

• In statistical applications, one is led to consider random estimates $\hat{G}^{(n)},\hat{F}_U^{(n)},\hat{H}^{(n)}$. Hence, if both the convergence $M_1(\hat{G}^{(n)},G)\vee M_1(\hat{F}_U^{(n)},F_U)\vee M_1(\hat{H}^{(n)},H)\to 0$ (L_1-consistency) and the boundedness of $\sigma^2_{\hat{G}^{(n)}}$ hold a.s. (respectively, in probability), then the results of the preceding theorem also hold a.s. (resp., in probability).

PROOF. Observe first that (28) immediately follows from (27) by virtue of standard properties of Wasserstein metrics. In order to prove (27), we construct a specific coupling of the laws $\mathcal{L}_{\hat{X}^{(T)}}$ and $\mathcal{L}_{X^{(T)}}$. Let $(V_k)_{n\in\mathbb{N}}$, $(V'_k)_{k\in\mathbb{N}}$ and $(V''_k)_{k\in\mathbb{N}}$ be three independent sequences of i.i.d. r.v.'s, uniformly distributed on $[0,1]$. For all $(n,k)\in\mathbb{N}^2$, set

$$\Delta T_k = G^{-1}(V_k), U_k = F_U^{-1}(V'_k), \theta_k = H^{-1}(V''_k),$$
$$\Delta\hat{T}_k^{(n)} = \hat{G}^{(n)^{-1}}(V_k), \hat{U}_k^{(n)} = \hat{F}_U^{(n)^{-1}}(V'_k), \hat{\theta}_k^{(n)} = \hat{H}^{(n)^{-1}}(V''_k),$$

and define recursively for $k\in\mathbb{N}$, $X_{k+1} = X_k e^{-\theta_k \Delta T_{k+1}} + U_{k+1}$ and $\hat{X}_{k+1}^{(n)} = \hat{X}_k^{(n)} e^{-\hat{\theta}_k^{(n)}\Delta\hat{T}_{k+1}^{(n)}} + \hat{U}_{k+1}^{(n)}$ with $X_0 = \hat{X}_0^{(n)} = x_0$, as well as $T_{k+1} = \Delta T_{k+1} + T_k$ and $\hat{T}_{k+1}^{(n)} = \Delta\hat{T}_{k+1}^{(n)} + \hat{T}_k^{(n)}$ with $T_0 = \hat{T}_0^{(n)} = 0$. For notational convenience, the superscript (n) is omitted in the sequel. Using in particular the fact that $x\geq 0 \mapsto e^{-x}$ is 1-Lipschitz, straightforward computations yield

$$\begin{aligned}
\left|\hat{X}_k - X_k\right| \leq x_0\{&\sum_{i=1}^{k}\theta_i\left|\Delta T_{i+1} - \Delta\hat{T}_{i+1}\right| + \sum_{i=1}^{k}\Delta\hat{T}_{i+1}\left|\theta_i - \hat{\theta}_i\right|\} \\
&+ \sum_{i=1}^{k} U_i\Big(\sum_{j=i}^{k-1}\theta_j\left|\Delta T_{j+1} - \Delta\hat{T}_{j+1}\right| + \sum_{j=i}^{k-1}\Delta\hat{T}_{j+1}\left|\theta_j - \hat{\theta}_j\right|\Big) \\
&+ \sum_{i=1}^{k}\left|\hat{U}_i - U_i\right|. \qquad (29)
\end{aligned}$$

Turning now to the coupling construction in continuous time, define $N(t) = \sum_{k\geq 1}\mathbb{I}_{\{T_k\leq t\}}$ and $\hat{N}(t) = \sum_{k\geq 1}\mathbb{I}_{\{\hat{T}_k\leq t\}}$, as well as $X(t) = X_{N(t)}exp(-\theta_{N(t)}(t - T_{N(t)}))$ and $\hat{X}(t) = \hat{X}_{\hat{N}(t)}exp(-\hat{\theta}_{\hat{N}(t)}(t-\hat{T}_{\hat{N}(t)}))$ for $t\geq 0$. Set also $T_k^+ = T_k\vee\hat{T}_k$ and $T_k^- = T_k\wedge\hat{T}_k$ for all $k\in\mathbb{N}$ and observe that

$$m_{\mathcal{H}}(\Gamma_{\hat{X}^{(T)}}, \Gamma_{X^{(T)}}) \leq \max_{0\leq k\leq N(T)\vee\hat{N}(T)} M_k, \qquad (30)$$

where

$$\begin{aligned}
M_k = &\sup_{T_k^+\leq t<T_{k+1}^-}\left|X(t-(\hat{T}_k - T_k)_+) - \hat{X}(t-(T_k-\hat{T}_k)_+)\right| + \left|T_k - \hat{T}_k\right| \\
&+ \sup_{T_{k+1}^-\leq t<T_{k+1}^+}\left|X(t-(t-T_{k+1})_+) - \hat{X}(t-(t-\hat{T}_{k+1})_+)\right| + \left|T_{k+1} - \hat{T}_{k+1}\right|,
\end{aligned}$$

denoting by $x_+ = 0\vee x$ the positive part of any $x\in\mathbb{R}$. And it follows from easy calculations that

$$M_k \leq \left|X_k - \hat{X}_k\right| + X_k(\Delta T_{k+1}\wedge\Delta\hat{T}_{k+1})\left|\theta_k - \hat{\theta}_k\right| + \sum_{i=1}^{k}\left|\Delta T_i - \Delta\hat{T}_i\right|$$

$$+ \left|X_{k+1} - \hat{X}_{k+1}\right| + \left|U_{k+1} - \hat{U}_{k+1}\right| + \sum_{i=1}^{k+1} \left|\Delta T_i - \Delta \hat{T}_i\right|.$$

By taking the expectation in (30) and then using the bounds $X_k \leq x_0 + \sum_{1 \leq i \leq k} U_i$ and (29) combined with Wald's lemma, straightforward computations yield

$$\begin{aligned}\mathbb{E}[m_{\mathcal{M}_2}^{(T)}(\hat{X}^{(T)}, X^{(T)})] \leq (1 + \mathbb{E}[N(T) \vee \hat{N}(T)])\{&2x_0(m_H M_1(G, \hat{G}) + m_{\hat{G}} M_1(H, \hat{H}) \\ &+ 3M_1(F_U, \hat{F}_U) + 2M_1(G, \hat{G}) + (x_0 + m_{F_U})(T + m_G + m_{\hat{G}}) \\ &\times M_1(H, \hat{H})\} + \mathbb{E}[(1 + N(T) \vee \hat{N}(T))^2] m_{F_U}(m_H M_1(G, \hat{G}) \\ &+ m_{\hat{G}} M_1(H, \hat{H})),\end{aligned}$$

denoting by m_F (resp. $m_{\hat{F}}$) the mean of the df F (resp. of the estimate $\hat{F}$), F being any of the df's G, F_U or H (notice that $m_{\hat{F}} \leq \nu_F + m_F$). Besides, there exist constants C, $C' < \infty$ s.t. $\mathbb{E}(N(T)) \vee \mathbb{E}(\hat{N}(T)) \leq CT$ and $\mathbb{E}(N(T)^2) \vee \mathbb{E}(\hat{N}(T)^2) \leq C'T^2$ (refer to Propositions 6.1 and 6.3 of chap. V in [3] for instance). Observe that the constants C and C' may be chosen independent from the integer n indexing the sequence $\hat{G}$, since by assumption the sequences $m_{\hat{G}}$ and $\sigma^2_{\hat{G}}$ are bounded. This establishes the desired result (27). ■

The next result now establishes the asymptotic validity of simulation estimators under general conditions.

Corollary 6 *Let (G, F_U, H) (resp. $(\hat{G}^{(n)}, \hat{F}_U^{(n)}, \hat{H}^{(n)})$ for $n \in \mathbb{N}$) be a triplet of df's on $\mathbb{R}_+$ defining a linear exposure process X (resp. $\hat{X}_{(n)}$) starting from $x_0 \geq 0$ and fulfilling the assumptions of Theorem 5. Let $0 < T \leq \infty$.*

• *Let $\mathcal{Q}$ be a measurable function mapping $D((0,T), \mathbb{R})$ into some metric space $(\mathcal{S}, D)$ with $\mathrm{Disc}(\mathcal{Q})$ as set of discontinuity points and such that $\mathbb{P}(X^{(T)} \in \mathrm{Disc}(\mathcal{Q})) = 0$. If $(\hat{G}^{(n)}, \hat{F}_U^{(n)}, \hat{H}^{(n)}) \to (G, F_U, H)$ in the L_1-Mallows distance, then we have the convergence in distribution*

$$\mathcal{Q}(\hat{X}_{(n)}^{(T)}) \Rightarrow \mathcal{Q}(X^{(T)}) \text{ in } (\mathcal{S}, D). \tag{31}$$

• *Suppose that G (respectively, $\hat{G}^{(n)}$) has finite variance σ_G^2 (resp. $\sigma^2_{\hat{G}^{(n)}}$) and H (respectively, $\hat{H}^{(n)}$) has finite mean. If $\sigma^2_{\hat{G}^{(n)}}$ remains bounded and $(\hat{G}^{(n)}, \hat{F}_U^{(n)}, \hat{H}^{(n)}) \to (G, F_U, H)$ in the L_1-Mallows distance, then for any Lipschitz function $\phi : (D((0,T), \mathbb{R}), m_{\mathcal{M}_2}^{(T)}) \to \mathbb{R}$, we have*

$$\mathbb{E}\left[\phi(\hat{X}_{(n)}^{(T)})\right] \to \mathbb{E}\left[\phi(X^{(T)})\right]. \tag{32}$$

PROOF. The first assertion derives from Theorem 5 and the convergence (in distribution) preservation result stated in Theorem 3.4.3 of [45], while the second one is an immediate consequence of the first assertion of Theorem 5 (see also [8]). ■

We conclude by giving several examples, illustrating how the results above apply to certain functionals of the exposure process in practice. Among the *time-dependent* and *steady-state* features of the exposure process, the following quantities are of considerable importance to practitioners in the field of risk assessment of chemicals in food and diet (see [36] and the references therein).

Mean exposure value. The mapping that assigns to any trajectory $X^{(T)} \in D((0,T),\mathbb{R})$ its mean value $T^{-1}\int_{t=0}^{T} X(t)dt$ is Lipschitz w.r.t. the distance $m_{\mathcal{M}_2}^{(T)}$. Hence, given consistent estimates $\hat{G}^{(n)}$, $\hat{F}_U^{(n)}$ and $\hat{H}^{(n)}$ of G, F_U and H, one may construct a consistent estimate of $\mathbb{E}[\int_{t=0}^{T} X(t)dt]$ by implementing a standard Monte-Carlo procedure for approximating the expectation $\mathbb{E}[\int_{t=0}^{T} \hat{X}_{(n)}(t)dt]$.

Maximum exposure value. In a similar fashion, the function $X^{(T)} \in D((0,T),\mathbb{R}) \mapsto \sup_{0\leq t\leq T} X(t)$ is Lipschitz w.r.t. the distance $m_{\mathcal{M}_2}^{(T)}$ (see Theorem 13.4.1 in [45] for instance) and under the assumptions of Theorem 5, the expected supremum $\mathbb{E}[\sup_{0\leq t\leq T} X(t)]$ is finite and may be consistently estimated by Monte-Carlo simulations.

First passage times. Given the starting point x_0 of the exposure process X, the distribution of the first passage time beyond a certain (possibly critical) level $x \geq 0$, *i.e.* the hitting time $\tau_x^+ = \inf\{t \geq 0, X(t) \geq x\}$, is also a characteristic of crucial interest for toxicologists. The mapping $X \in D((0,\infty),\mathbb{R}) \mapsto \tau_x^+$ being continuous w.r.t. the $\mathcal{M}_2$-topology (refer to Theorem 13.6.4 in [45]), we have $\hat{\tau}_x^+ = \inf\{t \geq 0, \hat{X}(t) \geq x\} \Rightarrow \tau_x^+$ as soon as $\hat{X} \Rightarrow X$.

Steady state mean exposure. In practice, one is also concerned with *steady-state* characteristics, describing the long term behavior of the exposure process. The steady-state mean exposure m_μ can be pertinently used as a quantitative indicator for chronic risk characterization. By virtue of Theorem 3 and Corollary 6, in an asymptotic framework stipulating that both $T \to \infty$ and $n \to \infty$, it can be consistently estimated by $\mathbb{E}[T^{-1}\int_{t=0}^{T} \hat{X}_{(n)}(t)dt]$ since one may naturally write

$$\begin{aligned}\mathbb{E}\left[T^{-1}\int_{t=0}^{T} \hat{X}_{(n)}(t)dt\right] - m_\mu &= \mathbb{E}\left[T^{-1}\int_{t=0}^{T} \hat{X}_{(n)}(t)dt\right] - \mathbb{E}\left[T^{-1}\int_{t=0}^{T} X(t)dt\right] \\ &+ \mathbb{E}\left[T^{-1}\int_{t=0}^{T} X(t)dt\right] - m_\mu.\end{aligned}$$

Besides, with regard to statistical applications, Theorem 5 paves the way for studying the asymptotic validity of bootstrap procedures in order to construct accurate confidence intervals (based on sample paths simulated from bootstrapped versions of the estimates $\widehat{G}^{(n)}, \widehat{F}_U^{(n)}$ and $\widehat{H}^{(n)}$). This is beyond the scope of the present paper but will be the subject of further investigation.

5 Application to methylmercury data

As an illustration of the toxicologic modeling presented above, some numerical results related to dietary methylmercury (MeHg) contamination are now exhibited. As previously mentioned, this chemical is present in sea food quasi-solely and a clear indication of its adverse effects on human heath has been given by observational epidemiological studies (see [44], [15] and [24] and references therein), leading recently regulatory authorities to develop sea food standards for protecting the safety of the consumer. Furthermore, the toxicokinetic variability in man of this substance has been thoroughly investigated in several studies (see [35], [40] and [39] for instance), almost all coming to the conclusion that the *half-life* of methylmercury in man (see remark 1) fluctuates around six weeks. In the present quantitative study, estimates of some of the level exposure indicators listed above are displayed for population subgroups: men, women and children aged under 15 (see Table 1 below).

The *half-life* $\log(2)/\theta$ of methylmercury is assumed to be distributed as a Gamma distribution with mean 6 weeks and standard deviation 3 days. For each group, an estimator of the intake distribution F_U has been computed from data collected through the national individual consumption survey INCA (see [14]) and contamination data released by French regulatory authorities, already used and described in [15] or [43]. For each consumption occurrence recorded in the INCA database, the MeHg intake is obtained as the cross product between sea food quantities consumed on a body weight scale (*i.e.*, divided by the body weight bw) and associated mean contamination levels. The histogram of these intakes (expressed in $\mu g/kg bw$) is used as an estimate of F_U. For each subgroup, a right censored Weibull model is fitted to estimate the distribution G of the inter-intake times from the INCA data.
The indicator estimates have been obtained by averaging over $M = 1000$ trajectory replications on $[0, T]$, with T equal to one year and $x_0 = 0$. Lower and upper bounds on indicators are given at a 95% confidence level. We observe in Table 1 that children are the most sensitive subgroup because they tend to eat more proportionally to their body weight even though their sea food consumption frequency is the lowest among the three groups.

Table 1: Comparison of the exposure process of men, women and children

	Men	Women	Children
Size of the data set for F_U	1605	1961	1612
Mean (in $\mu g/kgbw$)	0.206	0.230	0.487
Standard Deviation	0.274	0.352	0.768
Size of the data set for G	1769	2144	1816
Mean (in hours)	119	114	161
Standard Deviation	122	112	160
Mean (in $\mu g/kgbw$)	2.11	2.42	3.65
Lower bound	1.35	1.63	2.20
Upper bound	3.02	3.56	6.14
Maximum (in $\mu g/kgbw$)	4.14	4.81	7.74
Lower bound	2.54	2.88	4.18
Upper bound	6.56	8.40	16.55
First time passage in 5 when reached (in days)	226	196	150
Probability to reach 5 before 1 year	18%	32%	90%
First time passage in 10 when reached (in days)	>365	230	196
Probability to reach 10 before 1 year	0.0%	0.3%	16.6%
First time passage in 15 when reached (in days)	>365	>365	186
Probability to reach 15 before 1 year	0.0%	0.0%	3.6%

Figure 2 gives a time-plot of the mean exposure value for different values of T and x_0 in the men subgroup. The steady state mean exposure in the men subgroup can be seen to be approximatively equal to $2.5\mu g/kgbw$ and is nearly reached with a 1 year horizon. From a toxicological point of view, no "safe" steady state mean exposure upper value has been defined yet (proposing a dynamic stochastic model for the dietary contamination phenomenon is besides the main innovation of the present paper): the only quantity of reference is the so called *Provisional Tolerable Weekly Intake* (PTWI), which is considered to represent the contaminant dose an individual can ingest per week over all his life without appreciable risk. For methylmercury, the PTWI has been set to 1.6 $\mu g/kgbw$ by an international expert committee of FAO/WHO (see [21]). Hence, a deterministic exposure process of reference could be built by taking F_U as the Dirac mass on 1.6, G as the Dirac mass on 1 week and H as the Dirac mass on 6 weeks. This yields a steady state mean exposure of 14.6 $\mu g/kgbw$. From this angle, all studied subgroups remain far below this reference value.

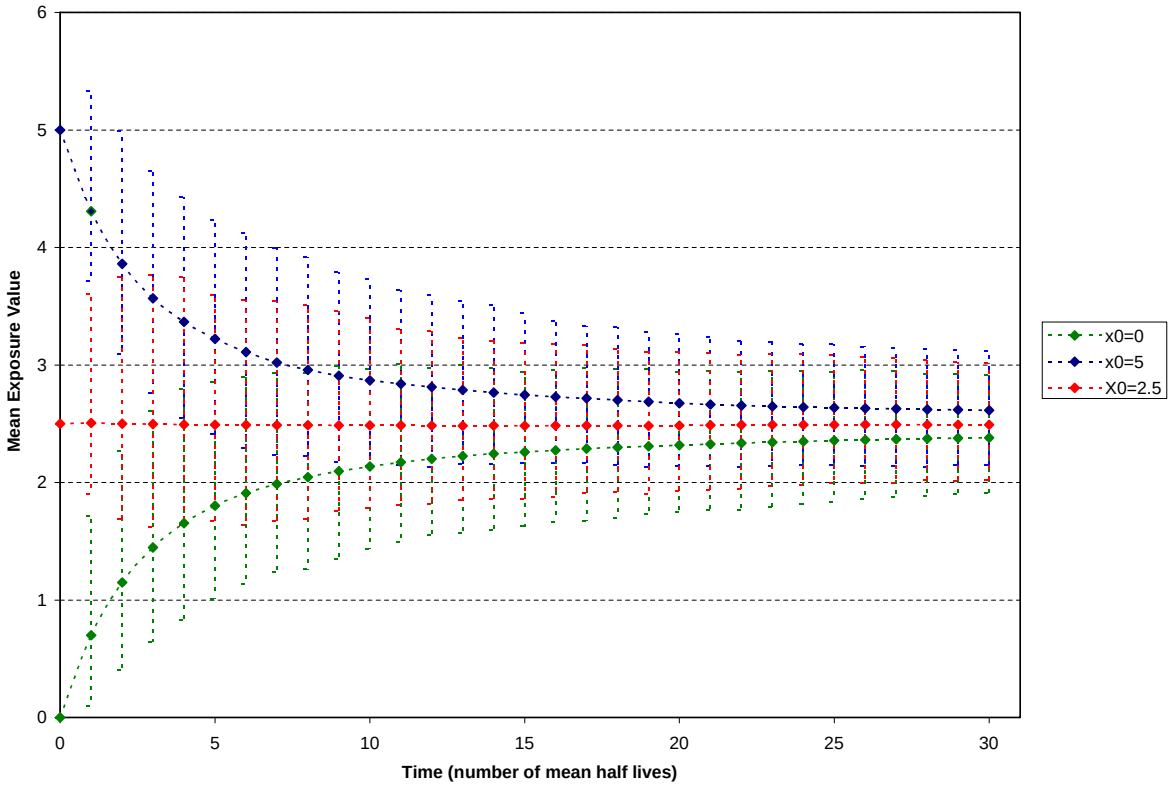

Figure 2: Mean Exposure Value as a function of time for the men subgroup

References

[1] B. Aberg, L. Ekman, R. Falk, U. Greitz, G. Persson, and J.O. Snihs. Metabolism of methyl mercury compounds in man. *Arch. Environ. Health*, 19(4):478–484, 1969.

[2] S. Asmussen. Subexponential asymptotics for stochastic processes: Extremal behavior, stationary distributions and first passage probabilities. *Adv. Appl. Probab.*, 8(2):354–374, 1998. Second Edition.

[3] S. Asmussen. *Applied Probability and Queues*. Springer-Verlag, New York, 2003.

[4] P. Bertail and S. Clémençon. Edgeworth expansions for suitably normalized sample mean statistics of atomic Markov chains. *Prob. Th. Rel. Fields*, 130(3):388–414, 2004.

[5] P. Bertail and S. Clémençon. Regenerative-block bootstrap for Markov chains. *Bernoulli*, 12(4), 2005.

[6] P. Bertail, M. Feinberg, J. Tressou, and P. Verger (Coordinateurs). *Analyse des Risques Alimentaires*. TEC&DOC, Lavoisier, Paris, 2006.

[7] P. Bertail and J. Tressou. Incomplete generalized U-Statistics for food risk assessment. *Biometrics*, 62(1):66–74, 2006.

[8] P.J. Bickel and D.A. Freedman. Some asymptotic theory for the bootstrap. *Ann. Statist.*, 9:1196–1217, 1981.

[9] P. Bougerol and N. Picard. Strict stationarity of generalized autoregressive processes. *Ann. Probab.*, 20(4):1714–1730, 1992.

[10] P.J. Brockwell, S.I. Resnick, and R.L. Tweedie. Storage processes with general release rule and additive inputs. *Adv. Appl. Probab.*, 14(2):392–433, 1982.

[11] S. Browne and K. Sigman. Work-modulated queues with applications to storage processes. *J. Appl. Probab.*, 29(3):699–712, 1992.

[12] O.L.V. Costa. Stationary distributions for piecewise-deterministic Markov processes. *J. Appl. Probab.*, 27(1):60–73, 1990.

[13] C. Cox, T.W. Clarkson, D.O. Marsh, L. Amin-Zaki, S. Tikriti, and G.G. Myers. Dose-response analysis of infants prenatally exposed to methylmercury: an application of a single compartment model to stringle-stranded hair analysis. *Environ. Res.*, 49:318–332, 1989.

[14] CREDOC-AFSSA-DGAL. *Enquête INCA (individuelle et nationale sur les consommations alimentaires)*. TEC&DOC, Lavoisier, Paris, 1999. (Coordinateur : J.L. Volatier).

[15] A. Crépet, J. Tressou, P. Verger, and J.Ch Leblanc. Management options to reduce exposure to methyl mercury through the consumption of fish and fishery products by the French population. *Regul. Toxicol. Pharmacol.*, 42:179–189, 2005.

[16] M.H.A. Davis. Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models. *J. R. Statist. Soc.*, 46(3):353–388, 1984.

[17] M.H.A. Davis. *Applied Stochastic Analysis*. Stochastics Monographs. Taylor & Francis, 1991.

[18] L. Edler, K. Poirier, M. Dourson, J. Kleiner, B. Mileson, H. Nordmann, A. Renwick, W. Slob, K. Walton, and G. Würtzen. Mathematical modelling and quantitative methods. *Food Chem. Toxicol.*, 40:283–326, 2002.

[19] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling Extremal Events for Insurance and Finance*. Applications of Mathematics. Springer-Verlag, 1999.

[20] R. Falk, J.O. Snihs, L. Ekman, U. Greitz, and B. Aberg. Whole-body measurements on the distribution of mercury-203 in humans after oral-intake of methylradiomercury nitrate. *Acta Radiol.*, 9:55–72, 1970.

[21] FAO/WHO. Evaluation of certain food additives and contaminants for methylmercury. Sixty first report of the Joint FAO/WHO Expert Committee on Food Additives, Technical Report Series, WHO, Geneva, Switzerland, 2003.

[22] W. Feller. *An Introduction to Probability and its Applications II*. Wiley, 1971. Second Edition.

[23] M.J. Gibney and H. van der Voet. Introduction to the Monte Carlo project and the approach to the validation of probabilistic models of dietary exposure to selected food chemicals. *Food Addit. Contam.*, 20(Suppl. 1):S1–S7, 2003.

[24] P. Grandjean, P. Weihe, R. White, F. Debes, S. Araki, K. Yokoyama, K.Murata, N. Sorensen, R. Dahl, and P. Jorgensen. Cognitive deficit in 7-year-old children with prenatal exposure to methylmercury. *Neurotoxicology Teratology*, 19:417–428, 1997.

[25] D.R. Grey. Regular variation in the tail behaviour of solutions of random difference equations. *Adv. Appl. Probab.*, 4(1):169–183, 1994.

[26] J. Jacod and A.N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer, New York, 1987.

[27] H. Kesten. Random difference equations and renewal theory for products of random matrices. *Acta Math.*, 131:207–248, 1973.

[28] T. Konstantopoulos and G. Last. On the use of Lyapunov function methods in renewal theory. *Stoch. Proc. App.*, 79:165–178, 1999.

[29] S.P. Meyn and R.L. Tweedie. A survey of Foster-Lyapunov conditions for general state space Markov processes. In *Proceedings of the Workshop on Stochastic Stability and Stochastic Stabilization*, Metz, France, 1994. Springer-Verlag.

[30] S.P. Meyn and R.L. Tweedie. *Markov Chains and Stochastic Stability*. Springer-Verlag, 1996.

[31] J.K. Miettienen. *Mercury, mercurials, and mercaptans*, chapter Absorption and elimination of dietary mercury and methyl mercury in man, pages 223–240. C.C. Thomas, Springfield, IL., M.W. Miller & T.W. Clarkson edition, 1973.

[32] S.T. Rachev. *Probability metrics and the stability of stochastic models*. Wiley, New York, 1991.

[33] S.T. Rachev and G. Rüschendorf. *Mass Transportation Problems I: Theory*. Springer, 1998.

[34] S.T. Rachev and G. Samorodnitsky. Limit laws for a stochastic process and random recursion arising in probabilistic modelling. *Adv. Appl. Probab.*, 27(1):185–202, 1995.

[35] T. Rahola, R.K. Aaron, and J. K. Miettienen. Half time studies of mercury and cadmium by whole body counting. In *Assessment of Radioactive Contamination in Man*, pages 553–562. Proceedings IAEA-SM-150-13, 1972.

[36] A. G. Renwick, S. M. Barlow, I. Hertz-Picciotto, A. R. Boobis, E. Dybing, L. Edler, G. Eisenbrand, J. B. Greig, J. Kleiner, J. Lambe, D.J.G. Mueller, S. Tuijtelaarsand P.A. Van den Brandt, R. Walker, and R. Kroes. Risk characterisation of chemicals in food and diet. *Food Chem. Toxicol.*, 41(9):1211–1271, 2003.

[37] G.O. Roberts and R.L.Tweedie. Bounds on regeneration times and convergence rates of Markov chains. *Stoch. Proc. Appl.*, 80:211–229, 1999.

[38] G.O. Roberts and R.L.Tweedie. Rates of convergence of stochastically monotone and continuous time Markov models. *J. Appl. Probab.*, 37(2):359–373, 2000.

[39] J. C. Smith and F. F. Farris. Methyl mercury pharmacokinetics in man: A reevaluation. *Toxicol. Appl. Pharmacol.*, 137:245–252, 1996.

[40] J.C. Smith, P.V. Allen, M.D. Turner, B. Most, H.L. Fisher, and L.L. Hall. The kinetics of intravenously administered methyl mercury in man. *Toxicol. Appl. Pharmacol.*, 128:251–256, 1994.

[41] J. Tressou. *Méthodes statistiques pour l'évaluation du risque alimentaire*. PhD thesis, Université Paris X, 2005.

[42] J. Tressou. Non parametric modelling of the left censorship of analytical data in food risk exposure assessment. *J. Amer. Stat. Assoc.*, 2006. In press.

[43] J. Tressou, A. Crépet, P. Bertail, M. H. Feinberg, and J. C. Leblanc. Probabilistic exposure assessment to food chemicals based on extreme value theory. application to heavy metals from fish and sea products. *Food Chem. Toxicol.*, 42(8):1349–1358, 2004.

[44] T. Tsubaki and K. Irukayama. *Minamata Disease: Methyl-mercury Poisoning in Minamata and Niigata, Japan.* ElsevierScientific, New York, 1977.

[45] W. Whitt. *Stochastic-Process Limits. An Introduction to Stochastic-Process Limits and their Application to Queues.* Springer, 2002.