MATHEMATICAL BIOSCIENCES AND ENGINEERING
http://www.mbejournal.org/
Volume 5, Number 1, January 2008 pp. 35–60

A STORAGE MODEL WITH RANDOM RELEASE RATE FOR MODELING EXPOSURE TO FOOD CONTAMINANTS

Patrice Bertail

Modal'X − Universite´ Paris X & CREST − LS
Universite´ Paris X, Ba^t. G, 200 avenue de la re´publique, 92001 Nanterre, FRANCE

Ste´phan Cle´menc¸on

LTCI − UMR 5141 Telecom Paris / CNRS & Me´t@risk − INRA
Telecom Paris − TSI, rue Barrault, 75634 Paris Cedex 13, FRANCE

Jessica Tressou

Me´t@risk − INRA & ISMT − HKUST
Hong Kong University of Science and Technology, ISMT, Clear Water Bay, HONG KONG

(Communicated by Susanne Ditlevsen)

Abstract. This paper presents the study of a 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 ..X, randomness of the coefficient .. 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 in describing the long−term behavior of the exposure process $(X_t)_{t=0}$ and assessing values such as the proportion of time the contaminant body burden is over a certain threshold. We also highlight the fact that the exposure process is generally not directly observable in practice and establish a validity framework for simulation−based statistical methods by coupling analysis. Eventually, applications to methyl mercury contamination data are considered.

1. Introduction. Certain foods may contain varying amounts of chemicals such as methyl mercury (present in seafood), dioxins (in poultry, meat), or mycotoxins (in cereals, dried fruits, etc.), which may cause major health problems when accumulated inside the body in excessive doses. Food safety is now a crucial public health concern in many countries (for 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

2000 Mathematics Subject Classification. Primary: 60J25; 62P10; Secondary: 92B15.
Key words and phrases. pharmacokinetics model, dietary contamination, food safety, piecewise−deterministic Markov process, simulation method, stability analysis.
The third author is supported by RGC grant #601906.

medicine, toxicology, and of course applied mathematics with the aim of developing rigorous methods for quantitative risk assessment. 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 [58], [24], [30], or [9]).

Static viewpoints for the probabilistic modeling of the quantity X of a given food contaminant ingested in a short period have been considered in recent works, mainly focusing on the tail behavior of X and allowing for computation of the probability that X exceeds a maximum tolerable dose (see [8], [57]). However, as highlighted in [62], such approaches for food risk analysis do not take into account the accumulating and eliminating processes occurring in the body, which naturally requires the introduction of time as a crucial description parameter of a comprehensive model (see also the discussion in [27]).

This paper proposes a dynamic modeling of exposure to a certain food contaminant, incorporating important features of the phenomenon, in a way that the model may account for the contaminant pharmacokinetics in the body following intakes. The case of methyl mercury food contamination shall serve as an example of the concepts and methods studied in this article: mathematical modeling of the pharmacokinetics behavior of methyl mercury (essentially present in seafoods) has received increasing attention in toxicology literature (see [45], [56], [55], [1], or [26]) and dose−response relationships have been extensively investigated for this contaminant, establishing clearly its negative impact on human health (refer to [15], [18]). 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. This 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 it allows the (content−dependent) release rate to be random, as strongly supported by biological modeling, and inter−intake times are not required to be exponentially distributed (the choice of a memoryless distribution being totally inadequate in the dietary context). Having practical use of the proposed exposure model for public health guidance in view (see [62]), we also discuss its relation to available data in the present paper: as sample paths of the exposure process cannot be observed in general, we set theoretical grounds for practical inference techniques based on intensive computer simulation methods. A thorough statistical analysis of toxicological and intake data based on the concepts and results developed in this article is carried out in a forthcoming companion paper (see [7]).

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), theoretical properties of the exposure process are thoroughly investigated in Section 3. Turning to the problem of estimating the

steady−state or time−dependent features of the model that are relevant from the toxicology viewpoint and taking into account the fact that the exposure process is not observable in practice, statistical procedures relying on simulation methods are presented and studied in Section 4. Finally, empirical studies related to methyl mercury 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. Technical proofs are postponed to the Appendix.

2. Modeling the exposure to a food contaminant. Suppose that an exhaus−tive 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. Regarding the chemical of interest, each type of food $p \in \{1, \ldots, P\}$ is con−taminated in random ratio $K^{(p)}$, with probability distribution $F_{K^{(p)}}$ on $\mathbb{R}_+$, the set of positive real numbers (we shall denote by $\mathbb{R}_+^*$ the set of strictly positive real numbers). Concerning this specific contaminant exposure, a meal may be viewed as a realization of a random variable (r.v.) $Q = (Q^{(1)}, \ldots, Q^{(P)})$ indicating the quantity of food of each type consumed, normalized by the body weight. For a meal Q drawn from a distribution F_Q on $\mathbb{R}_+^P$, cooked from foods of which toxicity is described by a contamination ratio vector $K = (K^{(1)}, \ldots, K^{(P)})$ drawn from the tensor product of distributions $F_K = \otimes_{p=1}^{P} F_{K^{(p)}}$, the global contaminant intake is

$$U = \sum_{p=1}^{P} K^{(p)} \cdot Q^{(p)} = \langle K, Q \rangle, \tag{1}$$

denoting by $\langle \cdot, \cdot \rangle$ the standard inner product on $\mathbb{R}^P$. Its probability distribution F_U is the image of $F_K \otimes F_Q$ by the inner product $\langle \cdot, \cdot \rangle$, assuming that the quan−tities 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 $d\lambda$, the Lebesgue measure on $\mathbb{R}_+$.

By convention, $T_0 = 0$ is chosen as time origin. The food contamination phe−nomenon through time for an individual of the population of interest may be clas−sically modeled by a marked point process $\{(T_n, Q_n, K_n)\}_{n=1}$ 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, $n = 1$ (refer to [21] for a recent account of the theory of point processes and of its numerous applications). The process $\{(T_n, Q_n)\}_{n=1}$ describing dietary behav−ior is assumed independent from the sequence $(K_n)_{n=1}$ 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, here we make the simplifying assumption that the marks Q_n, $n = 1$, form a sequence of independent and identi−cally distributed random variables (in the sequel we shall use the abbreviated form "i.i.d. r.v.'s"), with common distribution F_Q independent from the location times $(T_n)_{n=1}$. This assumption is acceptable for chemicals present in a few types of food, such as methyl mercury, our running example, but certainly not for all con−taminants. For chemicals present in many foods of everyday consumption such as Ochratoxin A (present in cereals, coffee, etc.), it would be necessary to introduce an 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 $...T_{n+1} = T_{n+1} - T_n$, $n = 1$, are i.i.d. with common probability distribution $G(dt) = g(t)dt$ and finite expectation $m_G = \int_{t=0}^{\infty} tG(dt) < \infty$, the sequence $(T_n)_{n=1}$ 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 $t = 0$. In between intakes, we assume that the contamination exposure process $X(t)$ is governed by the differential equation

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

denoting x's temporal derivative by $x..(t)$ and $..$ being a random parameter, taking its values in a set $T ... R^d$ with $d = 1$ say, and accounting in the modeling for fluctuations of the elimination rate due to metabolic factors at the intake times (the successive values $..n$, $n ... N$, of $..$ are thus fixed at times $T_0, T_1, . . .$, see Remark 2 below). The function $r(x, ..)$ is assumed to be strictly positive and continuous on $R_+^* \times T$, such that for all $.. ... T$, $r(0, ..) = 0$ (so that when $X(t)$ eventually reaches the level 0, the process stays at this level until the next intake) and for all $(, ..) ... (0, 1) \times T$:

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

Under these conditions, for any initial value $x(0) = 0$ and metabolic parameter value $.. ... T$, Eq. (2) clearly has a unique solution.

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.

Remark 1. (Pharmacokinetics modeling) In toxicology, Eq. (2) is widely used with $r(x, ..) = ..x$ for modeling the kinetics 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 [12], [54] or [29] for further details on linear pharmacokinetics models, also referred to as first−order kinetics models). In this context, the release parameter $\log 2/..$ is known as the biological half−life of the contaminant (the time required for X to decrease by half in the organism in absence of new contaminant intake). For methyl mercury (MeHg), our example in this paper, the half−life is known to fluctuate around six weeks (see [56] and the references therein). For such dietary contaminants, of which biological half−life is measured in weeks rather than days, it is naturally essential to take account of the kinetics for successful modeling of the exposure phenomenon.

Remark 2. (Modeling fluctuations in the metabolic rate) In a mathematical model for the evolution of food contaminant exposure through time, incorporating a certain amount of randomness in the elimination process due to possible metabolism changes may certainly contribute to make the modeling more plausible. In this first attempt, mainly motivated by the dietary methyl mercury case, we chose here to vary the metabolic parameter at each intake time, on grounds of parsimony: indeed, data collected generally consist of observed half−lives in a sample of individuals only (see [35, 28] for instance) and, to our knowledge, no quantitative study of the frequency of changes in metabolism (regarding chemical

elimination) has been carried out yet. However, our marked point process based model could be easily extended by considering competing risks between intake times and times when the metabolism changes. Another possible approach for modeling the fluctuations of .. could consist of using a stochastic differential equation (based on a geometric Brownian motion for instance).

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

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

Under these assumptions, we clearly have that $H(\{t_{..}(x_0, x) < 8\}) = 1$ for all $0 < x = x_0$. 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 next condition holds.

Condition (C_1): $H(\{t_{..}(x_0, 0) < 8\}) = 1$ for some $x_0 > 0$.

In such a case we would also have $H(\{t_{..}(x, 0) < 8\}) = 1$ for all $x = 0$. In this respect, it is noteworthy that, in the linear case mentioned in Remark 1, we have $t_{..}(x, 0) = 8$ for all .. > 0 and x > 0, meaning that the process cannot reach the level 0 in finite time (in contrast with pharmacokinetics models based on affine rates r(x) = a + b · x with a > 0). However, determining whether the chemical may be entirely removed from the body is purely a mathematical concern, due to existing limits of detection (LOD) inherent to analytical measurement techniques (see [33]).

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

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

with $U_n = hK_n, Q_n i$, n ... N, and $N(t) = \sum_{n...N} I_{\{T_n = t\}}$ as the number of intakes until time t, denoting by I_E the indicator function of any event E. The exposure process X is piecewise−deterministic with ca`d−la`g [1] trajectories (see a typical sample path in Fig. 1) and satisfies the equation

$$X(t) = X(0) + B(t) - \sum_{n=1}^{N(t)+1} \int_{T_{n-1}}^{T_n ...t} r(X(s), .._n)ds, \qquad (4)$$

X(0) denoting the total body burden in contaminant at initial time $T_0 = 0$ and with a ··· b = min(a, b) for all (a, b) ··· R^2. For an account of such piecewise deterministic processes, one may refer to [23] (see also [22] and ergodic results may be found in [17]).

For the continuous−time process thus defined to be markovian, one has to record the current value ..(t) = $\sum_{n...N} .._n I_{\{t...[T_n, T_{n+1}[\}}$ of the metabolic parameter as well as the backward recurrence time A(t) = $t - T_{N(t)}$ (the time since the last intake).

[1] Recall that any function x : R_+ ··· R is said ca`d−la`g if it is is everywhere right−continuous and has left limits everywhere: for all t > 0, $\lim_{s...t,\ s>t} x(s) = x(t)$ and $\lim_{s...t,\ s<t} x(s)$ exists and is finite.

Figure 1. Sample path of the exposure process X, modeling the evolution of the total body burden of a given dietary contaminant through time.

By construction, the process $(X(t), ..(t), A(t))_{t=0}$ is strongly markovian with in−finitesimal generator

$$Gf(x, .., t) = ..(t) \int_{u=0}^{\infty} \int_{..0...T} \{f(x + u, .., 0) - f(x, .., t)\} F_U(du) H(d.. 0) - r(x, ..)... _x f(x, .., t) + ..._t f(x, .., t), \quad (5)$$

denoting by $..(t) = g(t)/\int_{s=t}^{\infty} g(s)ds$ the hazard rate of the inter−intake times and provided that $f(., .., .) : (x, t) \mapsto f(x, .., t)$ is a bounded function with bounded continuous first derivatives in x and t for all $.. \cdots T$ (one may refer to [2] for an account of key notions of the theory of stochastic processes, oriented to biology applications).

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 $...T_1$ have the density $g_a(t) = g(a + t)/\int_{s=a}^{\infty} g(s)ds$, making the renewal process $(...T_n)_{n...N}$ possibly delayed, except 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 much more adequate (see section 5). Here and throughout we denote by $P_{x,a}$ the probability measure on the underlying space such that $(X(0), A(0)) = (x, a)$ and $..(0) \sim H$, and by $E_{x,a}[.]$ the $P_{x,a}$ −expectation for all $x = 0$ and a in supp(G), the support of the distribution G.

In the case when one neglects variability in the elimination process (i.e., when H is a Dirac measure) and under the additional assumption that the renewal times are exponentially distributed (making the process X itself markovian, which fa−cilitates the study but is not relevant to our application as emphasized above), this modeling boils down to a standard storage model with a general release rate (see [14] and [13] for instance). We refer to Chapter XIV in [4] for an account of such processes, widely used in operations research for modeling queuing/storage systems. Basic communication properties of the stochastic process $X = (X(t))_{t=0}$ may be established in a fashion very similar to the ones of the latter processes

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 [4]).

Theorem 2.1. Suppose that G(dx) = g(x)dx has infinite tail (so that arbitrarily long inter−intake times may happen with positive probability). Assume further that either g(x) > 0 on]0,] for some > 0 or else that F_U has infinite tail (i.e. with positive probability, either intake times may be arbitrarily close together or else intakes may be arbitrarily large). Then X reaches any state x > 0 in finite time with positive probability whatever the starting point, i.e. for all x_0 = 0, a ... supp(G), we have

$$P_{x_0,a}(t_x < 8) > 0, \tag{6}$$

with t_x = inf{t = 0 : X_t = x} as the (random) time needed for X to reach the level x. Furthermore, if condition (C_1) is fulfilled, then (6) still holds for x = 0. Besides, either X "goes to infinity" with probability one, i.e., is such that $P_{x_0,a}$({X(t) ... 8 , as t ... 8}) = 1 for all x_0 = 0, or X reaches any state x > 0 in finite time with probability one whatever the starting point, i.e., for all x_0 = 0, a ... supp(G),

$$P_{x_0,a}(t_x < 8\}) = 1. \tag{7}$$

If (C_1) is satisfied, then (7) 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 possibly critical threshold u > 0, $T^{-1}\int_0^T I_{\{X_t=u\}}\,dt$, for instance, are then asymptotically described by the distribution µ. Computing/estimating steady−state quantities would be then relevant for summarizing the exposure phenomenon in the long run and assessing the long−term toxicological risk. Beyond stochastic stability properties, determining the tail behavior of the steady−state distribution and evaluating the rate at which the exposure process converges to the stationary state is also of critical importance in practice. These questions are thoroughly investigated for linear pharmacokinetics models in the next section.

3. Probabilistic study in the 'linear kinetics' case. We now focus on the ergodicity properties of the exposure process X(t) in the specific case when, for a given metabolic state described by a real parameter .., the elimination rate is proportional to the total body burden in contaminant, i.e. r(x, ..) = ..x.

Here we suppose that T is a subset of 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 of numerous chemicals (see [29]). We shall show that studying the long−term behavior of X is reduced to investigating the properties of the embedded Markov chain $\tilde{X} = (X_n)_{n=1}$ that corresponds to the values taken by the exposure process just after intake times: $X_n = X(T_n)$ for all n = 1. By construction, the chain $\tilde{X}$ satisfies the following autoregressive equation with random coefficients

$$X_{n+1} = e^{-..\,T_{n+1}} X_n + U_{n+1}, \text{ for all } n = 1, \tag{8}$$

and has transition probability ..(x, dy) = p(x, y)dy with transition density

$$p(x, y) = \int_{.....T}^{8}\int_{t=1}^{} f_U(y - xe^{-..t})G(dt)H(d..), \tag{9}$$

for all (x, y) ... R_+^{*2}, where a ... b = max(a, b). Ergodicity of such real−valued Markov chains Y , defined through stochastic recurrence equations of the form $Y_{n+1} = a_n Y_n + ß_n$, where $\{(a_n, ß_n)\}_{n...N}$ 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 [25] for instance). Specialized to our setting, well known results related to such processes enable us to demonstrate that the embedded chains X~ is positive recurrent [2] under the assumption that $\log(1...U_1)$ has finite expectation (which is a very plausible hypothesis in the dietary context), as stated in the next theorem, and then to specify the tail behavior of the limiting probability distribution. Furthermore, the simple autoregressive form of Eq. (8) makes Foster−Lyapunov conditions easily verifiable for such Markov chains, in order to refine their stochastic stability analysis (we refer to [44] for an account of such key notions of the Markov chain theory).

Theorem 3.1. Under the assumptions of Theorem 2.1, the chain X~ is ..− irre− ducible [3]. Moreover, suppose that the following condition holds.

(H1): $E[\log(1 ... U_1)] < 8$.

Then X~ is positive recurrent with stationary probability distribution µ~.
If one assume further that f_U is continuous and strictly positive on R_+ and:

(H2): there exists some .. = 1 such that $E[U_1^{..}] < 8$,

then X~ is geometrically ergodic, µ~ has finite moment of order .. and there exist constants R < 8 and r > 1 such that, for all n = 1, x > 0,

$$\sup_{\{..,|..(z)|=1+z ..\}} \int_{y=0}^{8} ..(y).. \, {}^n(x, dy) − µ~(..) = R(1 + x ··)r^{−n}, \quad (10)$$

denoting by $..^n$ the n−th iterate of .. and with $µ~(..) = \int_{y=0}^{8} ..(y)µ~(dy)$ for any µ~− integrable function ...
Suppose finally that the condition (H1) and the next one simultaneously hold,

(H3): The r.v. U_1 is regularly varying with index .. > 0 (i.e. for all t > 0, $(1 − F_U(tx))/(1 − F_U(x)) ~ t^{−..}$ as x ... 8).

Then the stationary law µ~ has regularly varying tail with index ...

Remark 3. (Tail assumption for the intake distribution) The relevance of the regular variation assumption for modeling the tail behavior of dietary con− taminant intakes related to certain chemicals is strongly supported in [57] and [8]. In these works, various estimation strategies for tail distribution features such as the Pareto index .. involved in (H3) are also proposed and implemented on several food contamination and consumption data sets. We refer to [25] for an excellent ac− count of such notions arising in extreme values theory and techniques for modeling extremal events.

[2]Recall that a Markov chain $Y = (Y_n)_{n...N}$ with state space $(E, \mathcal{E})$ is positive recurrent if there exists a unique probability distribution µ on E that is invariant for its transition kernel .. (i.e $µ(dy) = \int_{x...E} µ(dx)..(x, dy)$), making then Y stationary (µ is then referred to as Y 's stationary distribution).

[3]A Markov chain $Y = (Y_n)_{n...N}$ with state space $(E, \mathcal{E})$ and transition ..(x, dy) is said ..− irreducible, .. being a s−finite measure on E, if, for all $A ... \mathcal{E}$ weighted by .., Y visits the subset A in finite time with positive probability whatever its starting point, i.e., $\sum_{n=1} ..(x, A) > 0$ for

As pointed out in [43], 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. Fortunately, given the explicit relationship between X and the embedded discrete−time chain X~ in our specific case, ergodicity of the continuous−time model and tail properties of its limiting distribution may be investigated based on the results established above for X~ under mild conditions. However, under more restrictive moment conditions for the inter−intake distribution, as the one stated below, a simple test function for which the generator (5) is shown to satisfy a geometric drift condition, may be nevertheless exhibited, so as to establish geometric ergodicity for the Markov process $\{(X(t), ..(t), A(t))\}_{t=0}$.

(H4): There exists $.. > 0$ such that $E[\exp(.....T_2)] < 8$.

Theorem 3.2. Under the assumptions of Theorem 2.1 and supposing that (H1) is fulfilled, X(t) has an absolutely continuous limiting probability distribution μ given by

$$\mu([u, 8[) = m^{-1} \int_{x=u}^{8} \int_{t=0}^{8} \int_{.....T}^{\log(x/u)} .. \, \mu\tilde{}(dx)G(dt)H(d..), \qquad (11)$$

in the sense [4] that $T^{-1} \int_0^T I_{\{X_t=u\}} \, dt \cdots \mu([0, u])$, $P_{x_0,a}$−a.s., as $t \cdots 8$ for all $x_0 = 0$ and $a \cdots \text{supp}(G)$. Furthermore,

- if (H3) holds and the set T is bounded, then μ is regularly varying with the same index as F_U,
- and if (H2) and (H4) hold, then $\{(X(t), ..(t), A(t))\}_{t=0}$ is geometrically recurrent. In particular, μ has finite moment of order .. and for all $(x, a) \cdots R_+^* \times \text{supp}(G)$, there exist constants $ß \ldots]0, 1[$, $B_a < 8$ such that

$$\sup_{..(z)=1+z} |E_{x,a}[..(X_t)] - \mu(..)| = B_a(1 + x \cdot\cdot)ß^t. \qquad (12)$$

Remark 4. (Tail behavior of the stationary distribution) When the U_n's are heavy−tailed, and under the assumption that the $...T_n$'s are exponentially distributed (making B(t) a time−homogeneous Le´vy process), the fact that the stationary distribution μ inherits its tail behavior from F_U has been established in [3] for deterministic release rates. Besides, when assuming G exponential and .. fixed, one may exactly identify the limit distribution μ in some specific cases (see section 8 in [13] or section 2 in Chap. XIV of [4]) 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.

Remark 5. (Practical relevance of steady−state features) Now that it is established that the exposure process settles to an equilibrium regime after a 'certain time', the question of specifying precisely what is meant by 'certain time' naturally arises. It would be pertinent to describe the long run behavior of the exposure process and assess the long−term toxicological risk by computing steady−state characteristics solely if the time necessary to reach the equilibrium state approximately may be considered as a reasonable horizon at the human scale. As an illustration, the amount of time needed to be roughly in steady−state from a

[4] We recall that a sequence of random variables $(X_n)_{n...N}$ is said to converge P−almost surely to a r.v. X when the event $\{\lim_{n...8} X_n = X\}$ happens with probability one, i.e., $P(\lim_{n...8} X_n = X) = 1$. One then standardly writes $X_n \ldots X$ P − a.s. in abbreviated form. More generally, any property holding true with probability one shall be said "holding true P−almost surely".

collection of datasets related to dietary MeHg contamination is evaluated through simulation in Section 5.

In order to exhibit connections between the exposure process $X = (X(t))_{t=0}$ 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 exposure 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 1. Under the assumptions of Theorem 2.1 and supposing that (H2) is fulfilled for $\alpha = 1$, we have for all $(x_0, a) \in \mathbb{R}_+ \times \mathrm{supp}(G)$

$$\bar{X}_T = \frac{1}{T}\int_{t=0}^{T} X(t)dt \to m_\mu, \quad P_{x_0,a}\text{-a.s.}, \tag{13}$$

as $T \to \infty$ with $m_\mu = \int_{x=0}^{\infty} x\mu(dx)$. Moreover, if (H2) is fulfilled with $\alpha = 2$, 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 $P_{x_0,a}$-distribution

$$\sqrt{T}(\bar{X}_T - m_\mu) \Rightarrow \mathcal{N}(0, \sigma^2) \text{ as } T \to \infty. \tag{14}$$

Remark 6. (Limiting variance of the sample mean exposure value) As will be shown in the proof below, the asymptotic variance σ^2 in (14) may be related to the limiting behavior of a certain additive functional of the Markov chain $((X_{\tau_n}, \tau_n, \ldots T_{\tau_{n+1}}))_{n=1}$. In [5] (see also [6]) 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 $E_\mu[X \mid X > u]$, denoting by $E_\mu[.]$ the expectation with respect to μ, after the fashion of risk evaluation in mathematical finance or insurance (see also [62]).

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. Performing measurements at all consumption times so as to record the food contamination levels is 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. In addition, 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 of various chemicals have been carried out. 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 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 distribution functions G, F_U and H. Hence, when no closed form analytic expression for the quantity of interest is 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 $L_{\hat{X}}$ corresponding to the estimated distribution functions $(\hat{G}, \hat{F}_U, \hat{H})$ and construct estimators based on the trajectories thus obtained.

Beyond stochastic modeling of the exposure phenomenon, the main goal of this paper is to provide theoretical grounds for the application of such statistical methods in toxicological risk evaluation. This leads up to investigate the stability of the stochastic model described in Section 2 with respect to G, F_U and H (stability analysis may be viewed as the counterpart of sensitivity analysis in a probabilistic framework, refer to [46] for an account of this topic), and consider the continuity problem consisting in evaluating a measure of closeness between L_X and $L_{\hat{X}}$ making the mapping $L_X \mapsto Q(X)$ continuous, Q being the functional of the trajectory of interest, a certain mapping defined on the Skorohod's space, i.e., the set of ca`d−la`g functions $x : \mathbb{R}_+ \to \mathbb{R}$. Hence, convergence preservation results may be obtained via the continuous−mapping approach as described in [63], 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 us to deal with the latter problem. For notational convenience, we omit indexing the probabilities and expectations considered in the sequel by $(x_0, 0)$.

Let $0 < T < \infty$. Considering the geometry of the (ca`d−la`g) sample paths of the exposure process X (see Fig. 1), we use the M_2 topology on the Skorohod's space $D_T = D([0, T], \mathbb{R})$ induced by the Hausdorff distance on the space of completed graphs (the completed graph of $x \in D_T$ 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 J_2 topology would actually be sufficient for our purpose, refer to [36] or [63] 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 $L_{X^{(T)}}$ and $L_{\hat{X}^{(T)}}$ 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_T is defined as follows (refer to [47], [10]):

$$W_1^{(T)}(L, L_0) = \inf_{Z_0 \sim L_0,\, Z \sim L} E[m_{M_2}^{(T)}(Z_0, Z)],$$

where the infimum is taken over all pairs (Z_0, Z) with marginals L_0 and L and $m_{M_2}^{(T)}(Z_0, Z) = m_H^{(T)}(G_{Z_0}, G_Z)$, denoting by G_{Z_0} and G_Z the completed graphs of Z_0 and Z and by $m_H^{(T)}$ 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 [10]).

As claimed in the next theorem, the law $L_{\hat{X}^{(T)}}$ gets closer and closer to $L_{X^{(T)}}$ as the distribution functions $\hat{G}$, $\hat{F}_U$ and $\hat{H}$ respectively tend to G, F_U and H in the Mallows sense. For $p \in [1, \infty)$, we denote by $M_p(F_1, F_2) = \left(\int_0^1 |F_1^{-1}(t) - F_2^{-1}(t)|^p\, dt\right)^{1/p}$

the L_p−Mallows distance between two distribution functions F_1 and F_2 on the real line.

Theorem 4.1. Let (G, F_U, H) (respectively, $(G^{(n)}, F_U^{(n)}, H^{(n)})$ for $n \in \mathbb{N}$) be a triplet of distribution functions on $\mathbb{R}_+$ defining a linear exposure process X (respectively, $X^{(n)}$) starting from $x_0 = 0$ and fulfilling Theorem 2.1's assumptions and (H2) with $\mu = 1$. Suppose that $(G^{(n)}, F_U^{(n)}, H^{(n)})$ tends to (G, F_U, H) in the L_1−Mallows distance as $n \to \infty$. Assume further that G (respectively, $G^{(n)}$) has finite variance σ_G^2 (resp., $\sigma_{G^{(n)}}^2$) and H (resp., $H^{(n)}$) has finite mean, m_H. If $\sigma_{G^{(n)}}^2$ remains bounded, then:

$$\sup_{T>0} T^{-2} W_1^{(T)}(L_{X^{(T)}}, L_{X^{(T)}}) \to 0, \text{ as } n \to \infty. \tag{15}$$

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

$$X_{(n)}^{(T)} \Rightarrow X^{(T)} \text{ in } D_T, \text{ as } n \to \infty. \tag{16}$$

Remark 7. Before showing how this theoretical result applies to the problem of approximating/estimating general functionals of the exposure process, a few remarks are in order.

- We point out that similar results hold for the L_p−Wasserstein Kantorovich distance with $p \in [1, \infty)$ under suitable moment conditions.
- It may also be convenient to consider the function space $D_\infty = D([0, \infty), \mathbb{R})$ in which X has its sample paths and on which the metric

$$m_{M_2}^{(\infty)}(x, x^0) = \int_{t=0}^{\infty} 2^{-t} m_{M_2}^{(t)}(\{x(s)\}_{s\in[0,t]}, \{x^0(s)\}_{s\in[0,t]})dt$$

  for $(x, x^0) \in D_\infty^2$ may be considered. It is noteworthy that (15) also immediately provides a control of the L_1−Wasserstein distance $W_1^{(\infty)}$ corresponding to that metric between L_X and $L_{X^{(n)}}$.
- In statistical applications, one is led to consider random estimates $G^{(n)}$, $F_U^{(n)}$, and $H^{(n)}$. Clearly, if both the convergence $(G^{(n)}, F_U^{(n)}, H^{(n)}) \to (G, F_U, H)$ (i.e., 'L_1−consistency' of the distribution estimates) and the boundedness of $\sigma_{G^{(n)}}^2$ holds almost surely, then the results of the preceding theorem (and those stated in the next corollary) also hold almost surely.

By demonstrating that good approximations/estimations of the distributions H, G and F_U also induces good approximation/estimation of general functionals of the exposure process, the next result establishes the asymptotic validity of simulation estimators under general conditions. In general, provided that the instrumental distribution estimates at our disposal are accurate, we may thus treat simulated sample paths as if they were really exposure trajectories of individuals in the population of interest.

Corollary 1. Let (G, F_U, H) (respectively $(G^{(n)}, F_U^{(n)}, H^{(n)})$ for $n \in \mathbb{N}$) be a triplet of distribution functions on $\mathbb{R}_+$ defining a linear exposure process X (respectively $X^{(n)}$) starting from $x_0 = 0$ and fulfilling the assumptions of Theorem 4.1. Let $0 < T < \infty$.

(i) Let Q be a measurable function mapping D_T into some metric space (S, D) with $\mathrm{Disc}(Q)$ as set of discontinuity points and such that $P(X^{(T)} \in \mathrm{Disc}(Q)) = 0$. Then, $Q(X_{(n)}^{(T)}) \Rightarrow Q(X^{(T)})$ as $n \to \infty$.

0. Then we have the convergence in distribution

$$Q(\hat{X}_{(n)}^{(T)}) \dots Q(X^{(T)}) \text{ in } (S, D).$$

(ii) For any Lipschitz function $f : (D_T, m_{M_2}^{(T)}) \cdots R$, we have

$$E^h f(\hat{X}_{(n)}^{(T)})^i \dots E^h f(X^{(T)})^i \text{ as } n \dots 8.$$

Proof. The first assertion derives from Theorem 4.1 and the convergence (in distribution) preservation result stated in Theorem 3.4.3 of [63], while the second is an immediate consequence of the first assertion of Theorem 4.1 (see also [10]).

We conclude this section by giving several examples, illustrating how the results above apply to certain functionals of the exposure process in practice. Among the time−dependent 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 [50] and the references therein).

Maximum exposure value. The mapping that assigns to any finite length trajectory $\{x(t)\}_{0=t=T} \cdots D_T$ its maximum value $\sup_{0=t=T} x(t)$ is Lipschitz with respect to the Hausdorff distance $m_{M_2}^{(T)}$ (see Theorem 13.4.1 in [63]). Under the assumptions of Theorem 4.1, the expected supremum is finite and, given consistent estimates $G^{\hat{}}(n)$, $F_U^{\hat{}(n)}$ and $H^{\hat{}}(n)$ of G, F_U and H, one may thus construct a consistent estimate of $E[\sup_{0=t=T} X(t)]$ by implementing a standard Monte−Carlo procedure for approximating the expectation $E[\sup_{0=t=T} \hat{X}_{(n)}(t)dt]$.

Mean exposure value. The function $\{x(t)\}_{t...[0,T]} \cdots D_T 7... T^{-1} R_{t=0}^T x(t)dt$ being continuous with respect to M_2−topology, we have

$$T^{-1} \int_{t=0}^T \hat{X}^{(n)}(t)dt \dots T^{-1} \int_{t=0}^T X(t)dt$$

as soon as $\hat{X}_{(n)}^{(T)} \dots X^{(T)}$ in D_T. By straightforward uniform integrability arguments, it may be seen that convergence in mean also holds, so that the mean exposure value $E[T^{-1} R_{t=0}^T X(t)dt]$ may be consistently estimated by Monte−Carlo simulations.

Average time spent over a critical threshold. Let $u > 0$ be some critical level. In a very similar fashion, it follows from the continuity of $\{x(t)\}_{t...[0,T]} \cdots D_T 7... T^{-1} R_{t=0}^T I_{\{x(t)=u\}} dt$, that a Monte−Carlo procedure also allows to estimate the expectation of the average time spent by the exposure process above the threshold value u, namely $E[T^{-1} R_{t=0}^T I_{\{X(t)=u\}} dt]$.

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 $x = 0$, i.e., the hitting time $t_x^+ = \inf\{t = 0, X(t) = x\}$, is also a characteristic of crucial interest for toxicologists. The mapping $X \dots D((0, 8), R) 7... t_x^+$ being continuous w.r.t. the M_2−topology (refer to Theorem 13.6.4 in [63]), we have $\hat{t}_x^+ = \inf\{t = 0, \hat{X}(t) = x\} \dots t_x^+$ as soon as $\hat{X} \dots X$.

In practice, one is also concerned with steady−state characteristics, describing the long term behavior of the exposure process. For instance, the steady−state mean exposure $m_{..}$, or the limiting time average spent above a given critical value $u > 0$,

$\mu([u, 8[) = \lim_{T \to 8} T^{-1} \int_{t=0}^{T} I_{X(t)=u}\, dt$, can be pertinently used as quantitative indicators for chronic risk characterization (see also [62]). As shall be seen below, such quantities may be consistently estimated in a specific asymptotic framework stipulating that both T and n tend to infinity. As a matter of fact, one may naturally write

$$E\left[\left(\frac{1}{T}\int_{t=0}^{T} \hat{X}^{(n)}(t)dt - m_\mu\right)^2\right] = \left(E\left[\frac{1}{T}\int_{t=0}^{T} \hat{X}^{(n)}(t)dt - \frac{1}{T}\int_{t=0}^{T} X(t)dt\right]\right)^2 + \left(E\left[\frac{1}{T}\int_{t=0}^{T} X(t)dt - m_\mu\right]\right)^2. \quad (17)$$

The last term between brackets on the left hand side of (17) tends to 0 as $T \to 8$ by virtue of Theorem 3.2, while it follows from the coupling argument of Theorem 4.1 (see A4 in the Appendix) that the first term on the right hand side is less than $T \times M_1(\hat{F}_U^{(n)}, F_U) + T^2 \times (M_1(\hat{G}^{(n)}, G) + M_1(\hat{H}^{(n)}, H))$ up to a multiplicative constant. Hence, if T and n simultaneously tend to infinity in a way that the latter quantity converges to 0, consistency of $E\left[T^{-1}\int_{t=0}^{T} \hat{X}^{(n)}(t)dt\right]$ as an estimator of m_μ is clearly established.

In addition, with regard to statistical applications, Theorem 4.1 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 $\hat{G}^{(n)}$, $\hat{F}_U^{(n)}$ and $\hat{H}^{(n)}$). This is beyond the scope of the present paper but will be the subject of future investigation.

5. Application to methyl mercury data. As an illustration of the mathemat−ical toxicological model analyzed above, some numerical results related to dietary methyl mercury (MeHg) contamination are now exhibited. As previously men−tioned, this chemical is present in seafoods quasi−solely and a clear indication of its adverse effects on human heath has been given by observational epidemiologi−cal studies (see [60], [15] [20], and [31] and references therein), leading regulatory authorities to recently develop seafood standards for protecting the safety of the consumer.

At present, dietary risk assessment is conducted from a static viewpoint, com−paring the weekly intakes to a reference dose called Provisional Tolerable Weekly Intake (PTWI), which is considered to represent the contaminant dose an indi−vidual can ingest each week in his lifetime without appreciable risk. For methyl mercury, the PTWI has been set to 1.6 micrograms per kilogram of body weight per week (µg/kgbw/w in abbreviated form) by the international expert commit−tee of FAO/WHO (see [28]). Another reference dose of 0.7 µg/kgbw/w, estab−lished from a previous evaluation by the (U.S.) National Research Council [61], is sometimes used for a more conservative perspective. Hence, in a dynamic ap−proach, a deterministic exposure process of reference for risk assessment could be built by considering weekly intakes exactly equal to one of these static reference doses (d = 1.6 or 0.7) and a fixed mean half−life HL expressed in weeks. In this case, the body burden at the n^{th} intake is given by the (affine) recurrence relation $X_n = \exp(-\log(2)/HL \times 1)X_{n-1} + d$. The dynamic reference dose is obtained by taking the limit as n tends to infinity: $X_{ref,d} = d/(1 - 2^{-1/HL})$. Numerically, this yields $X_{ref,0.7} = 6.42$ µg/kgbw and $X_{ref,1.6} = 14.67$ µg/kgbw when MeHg biological

half−life is fixed to six weeks as estimated in [56].

Datasets and empirical estimates of distributions F_U, G and H. Food contamination data related to fish and other seafoods available on the French market have been collected by accredited laboratories from official national surveys performed between 1994 and 2003 by the French Ministry of Agriculture and Fisheries [40] and the French Research Institute for Exploitation of the Sea [34]. Our dataset is comprised of 2832 analytical data.

The national individual consumption survey INCA (see [19]) provides the quantity consumed of an extensive list of foods over a week, including fish and seafoods, as well as the time when consumption occurred with the information about the type of meal (breakfast, lunch, dinner or snack). The survey is composed of two samples: 1985 adults (15 years and older), and 1018 children (3 to 14 years). However, as shown by the hazard characterization step (see [28]), the group that is the most critically exposed to neuro−developmental adverse effects of MeHg are foetus: the MeHg in the mother's fish diet can pass into the developing foetus and cause irreversible brain damage. Here we thus focus on women of childbearing age (between 15 and 45).

For simplicity, MeHg intakes are computed at each observed meal through a deterministic procedure currently used in national and international risk assessments. From the INCA food list, 92 different fish or seafood species are determined and a mean level of contamination is computed from the contamination data, as in [20, 59]. Intakes are then obtained by applying relation (1). For comparability sake, all consumptions are divided by the associated individual body weight, also provided in the INCA survey.

(a) Intake distribution (b) Inter−intake time distribution

Figure 2. Probability plots for the distribution fitting (Adult Female, 15−45).

After the work of [57], the intake distribution F_U is modeled by a heavy−tailed distribution, namely the Burr distribution (of which cumulative distribution function is of the form $(1 - (1 + x^c)^{-k})$, with $c > 0$ and $k > 0$). It is noteworthy that it fulfills assumption H3 (see Remark 3). It is fitted here by means of standard maximum likelihood techniques (see Fig. 2(a) below for a probability plot illustrating the goodness of the fit).

The times of consumption available from the INCA survey allow us to compute the inter−intake times or at least produce right censored values (providing then the information that some durations between successive intakes are larger than a certain time). A Gamma distribution (which has increasing hazard rate) is retained for modeling the inter−intake distribution G: its parameters are fitted using a right censored maximum likelihood procedure. As shown by the probability plot displayed in Fig. 2(b), the chosen distribution (Gamma) provides a good fit for the left (uncensored) part of the distribution.

The pharmacokinetics of MeHg has been thoroughly investigated in several studies (see [49], [55], [56], and [35] for instance), almost all coming to the conclusion that the half−life of methyl mercury in man (see Remark 1) fluctuates around six weeks. As we could not dispose of any raw data related to MeHg half−life in the human body, based on the most documented study (in which collected half−life data (in days) are indicated to range from 36 to 65 and correspond to a sample mean of 44, refer to [56] and [35]), a Gamma distribution with mean 44 days and 5th percentile 36 days is chosen here for modeling the variability of the biological half−life log 2/...

Table 1 sums up the characteristics of the three input distributions, with the convention that a Gamma distribution with parameter a and ß has density given by $G(a)^{-1} ß^{-a} x^{a-1} \exp(-x/ß)$, and mean aß. Recall that the tail index of the Burr distribution is given by .. = $(ck)^{-1}$ and its moment of order r is finite if ck > r, it is then given by G(k − r/c) × G(1 + r/c)/G(k).

Table 1. Parameters of the input distributions (Adult Female, 15−45; fitted by maximum likelihood).

Intake distribution	F_U	(Unit: µg/kgbw/meal)
Sample size	n	1088
Burr Parameters	c	0.95
	k	4.93
Tail index	..	0.214
Mean Intake	m_{F_U}	0.243
Inter−intake time distribution	G	(Unit: hour)
Sample size	n	1214
Proportion of censored data		47.4%
Gamma Parameters	a	1.07
	ß	117.21
Mean inter−intake time	m_G	125
Half−Life distribution		(Unit: hour)
Gamma Parameters	a	13.6
	ß	77.4
Mean half−life		44 × 24

Time to steady−state. As underlined in Remark 5, the question of determining how much time is needed for the exposure process to be roughly at equilibrium is crucial for assessing the practical relevance of steady−state characteristics. In order to evaluate the time to steady−state, Monte−Carlo simulations have been carried out: using the instrumental distributions described above, $M = 1000$ trajectories

have been started from different initial values x_0. For each path, the (temporal) mean exposure over the time interval $[0, T]$ is computed and individual results have been averaged. In Figure 3(a), the resulting estimates for the Adult Female (15−45) subpopulation are displayed, as time T grows: as expected, all empirical mean exposures converge to the same quantity (namely m_μ) and the relative error is lower than 10% after 29 half−lives (3.5 years approximately), whatever the starting value x_0.

The same procedure is used for the mean time spent beyond the reference threshold $u = X_{\mathrm{ref},0.7}$: as shown by Figure 3(b), the limit is approximately reached after 70 half−lives (about 8.5 years), whatever the initial state x_0. The convergence is slower in this case, since the quantity of interest is related to an event of relatively small probability (see also Remark 10).

Naturally, the time to steady−state strongly depends on the initial value x_0 and of the functional of interest. However, on the basis of these simulation results, it may be stated that, for realistic initial values, the time to steady−state for basic quantities as the ones considered here fluctuates between 3 and 10 years, which is a reasonable horizon at the human scale.

(a) Mean exposure versus time

(b) Average time spent over u versus time ($u = X_{\mathrm{ref},0.7} = 6.42$ µg/kgbw)

Figure 3. Convergence to steady−state (Adult Female, 15−45; $x_0 = 0, 1.2, 2.4, 3.6, 4.8, 6$).

Remark 8. (Convergence rate bounds) The problem of determining computable bounds on the convergence rate of ergodic Markov processes has recently received much attention in the applied probability literature (see [42], [53], [39], [51] or [52]). Using suitably calibrated parameters of the drift and minorization conditions (Equations (18), (19), (23), and (22)) established along the proofs of theorems 3.1 and 3.2 in the Appendix, rough numerical estimates of the constants involved in rate bounds (10) and (12) can be computed from Theorem 2.2 in [52] (in the discrete−time case) and Theorem 4.1 in [52] (in the continuous−time setup). Explicit computations based on these theoretical results shall be carried out in a forthcoming paper (see [7]). However, as such computable bounds are quite loose in general and are related to the total variation norm (whereas one rather focuses on specific functional of interest in practice), the problem is handled here by exploiting the simulation approach.

Estimation by computer simulation. We now focus on estimating certain relevant time−dependent and steady−state quantities among those enumerated in the previous section via the simulation approach studied in Section 4 from our MeHg datasets. Estimates of the quantities of interest are computed by averaging over $M = 1000$ replicated trajectories on $[0, T]$, taking T equal to 5 years. In order to ensure that the retained trajectories are approximately stationary, a burn−in period of 5 years is used (see the preceding paragraph). Numerical results related to the estimation of the steady−state mean exposure (m_μ), the probability to exceed the dynamic reference doses in steady−state ($\mu([X_{ref,0.7}, \infty[)$ and $\mu([X_{ref,1.6}, \infty[)$, the mean time to run over the lowest reference dose ($E_\mu[t_{X_{ref,0.7}}]$) and the expected maximum exposure over 5 and 10 years ($E_\mu[\max_{t=5/10 \text{ years}} X(t)]$) are displayed in Table 2.

Table 2. Estimated features of the exposure process (Adult Female, 15−45).

Parameter	Unit	Estimate
m_μ	(µg/kgbw)	2.92
$\mu([X_{ref,0.7}, \infty[)$	(%)	0.575
$\mu([X_{ref,1.6}, \infty[)$	(%)	0.003
$E_\mu[\max_{t=5 \text{ years}} X(t)]$	(µg/kgbw)	6.63
$E_\mu[\max_{t=5 \text{ years}} X(t)]$	(µg/kgbw)	7.41

We observe that the average time spent over the EU−based threshold ($X_{ref,1.6}$) or the US−based one ($X_{ref,0.7}$) are close to zero in the Adult Female (15−45) subpopulation, resp. 0.003% and 0.575%. Regarding the time required to reach such threshold levels, further simulations have been conducted using the estimated stationary mean as the initial point (namely, $x_0 = 2.92$). Only the distribution of the time to reach the US−based threshold ($X_{ref,0.7}$) has been estimated using a standard Monte−Carlo procedure As explained in Remark 10 below, estimating the distribution of the time required to run over level $X_{ref,1.6}$ involves computing rare event probabilities and thus requires the use of more sophisticated simulation methods. Over $M = 1000$ trajectories, the mean (respectively, the median) of $t^+_{X_{ref,0.7}}$ is 7.23 years (resp. 5.05 years). Figure 4 displays the (highly skewed) Monte−Carlo distribution estimate (obtained by a kernel estimation built over the $M = 1000$ simulated values using a standard procedure) of the time to run beyond $X_{ref,0.7}$ for the studied subpopulation.

Remark 9. (Sensitivity to the instrumental distributions) The distribution models used here for the governing probability measures F_U and G have been chosen because they provide a good overall fit to the data (and may be easily seen to satisfy all the assumptions required in the ergodicity and stability analyses). According to our own practical experience, the numerical results displayed here would not have been significantly different, if one had chosen to model G by a Weibull distribution for instance. However, in a future study (see [7]) special attention shall be given to the statistical issues of validating the mathematical toxicological model, by investigating how sensitive the latter is to changes with respect to the estimation method chosen (considering also the use of nonparametric approaches).

Figure 4. Monte−Carlo distribution for the time to run over u for the Adult Female (15−45) subpopulation ($u = X_{ref,0.7} = 6.42$ µg/kgbw, with $x_0 = 2.92$).

Remark 10. (Naive Monte−Carlo simulation and high threshold) From the perspective of public health guidance practice, it is of prime importance to evaluate the probability of occurrence of rare (extremely risky) events, the probability to exceed a large threshold u such as $u = X_{ref,1.6}$ for instance. In this respect, we point out that the naive Monte−Carlo simulation proposed here leads to estimate this probability by zero (see Table 2), so seldom this threshold is reached on a time interval [0, T] of reasonable length. Treading in the steps of [16], it is shown in [7] that one may remedy to this problem by implementing a suitable particle filtering algorithm.

Appendix A. Technical proofs.

A.1. Proof of Theorem 3.1. From conditions required by Theorem 2.1, aperiodicity and irreducibility properties are immediately established for the discrete−time chain X~. In addition, under mild irreducibility conditions, the stability of the random coefficients autoregressive model on R^d

$$Y_{n+1} = a_n Y_n + ß_n,$$

where $(a_n, ß_n)$, $n = 1, \ldots$ are i.i.d. r.v.'s on $R_+^* \times R^d$, has been investigated in detail since the seminal contribution of [37] (see [48] and the references therein). Under the assumption that $E[\log(1 \ldots kß_1k)] < 8$ and $E[\log(1 \ldots a_1)] < 8$, it is well known that a sufficient and necessary condition for the chain X to have a (unique) probability measure is that $E[\log(a_1)] < 0$ (see Corollary 2.7 in [11] for instance). Based on this result, it is then straightforward that, under the assumptions of Theorem 2.1 and (H1), the chain 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 $(Y_n)_{n \ldots N}$ may be carried out by establishing suitable conditions for the 'drift' $\ldots V(y) = E[V(Y_1) \mid Y_0 = y] - V(y)$ 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 [44]) toward which the chain drifts in the sense that:

$$...V(x) = -f(x) + bI_{\{x...S\}}, \tag{18}$$

for some 'norm−like' function $f(x) = 1$ and $b < 8$. Now for the chain $X\tilde{}$, any compact interval $[0, s]$ with $s > 0$ is small. Indeed, it follows from (9) that for all $x ... [0, s]$, the minorization condition below holds:

$$..(x, .) = d_S \cdot U_S(.), \tag{19}$$

with $d_S = s \times \inf_{y...[0,s]} f_U(y) > 0$ and denoting by $U_S(.)$ the uniform probability distribution on $[0, s]$. When $.. = 1$ for instance, take $V(x) = 1 + x$. The affine drift related to $X\tilde{}$ is given by

$$...V(x) = -cx + E[U_1],$$

with $c = 1 - E[e^{-..1...T_2}] > 0$. Choosing $S = [0, s]$ with $s = 1 + 2E[U_1/c]$, (18) is fulfilled with $f(x) = cV(x)/2$ and $b = E[U_1] + c/2$. Applying Theorem 15.0.1 in [44], we thus get that $X\tilde{}$ is geometrically ergodic with invariant probability measure $\mu\tilde{}$ such that $\mu\tilde{}(V) = \int_{x=0}^{8} V(x)\mu\tilde{}(dx) < 8$. In particular, $\mu\tilde{}$ has finite expectation and there exist constants $r > 1$, $R < 8$ such that for all $x > 0$:

$$\sum_{n=0}^{8} r^n k.. P^n(x, .) - \mu\tilde{}k_V = RV(x), \tag{20}$$

with $k..k_V = \sup_{..:|..|=V} \int_R ..(x)..(dx)$ for all bounded measure $..$ on the real line. When $V = 1$, $k.k_V$ coincides with the total variation norm $k.k_{TV}$. For $.. > 1$, the results is proved in a similar fashion by taking $V(x) = 1 + x^{..}$.

Finally, the last assertion of Theorem 3.1 immediately derives from Theorem 1 in [32].

A.2. Proof of Theorem 3.2. Set $X_0 = X(0)$. Observe that for all $t > 0$, $X(t) = X_{N(t)} e^{-..A(t)}$, so that $X(t) = X_{N(t)}$. Hence, we naturally have $\{X(t) ... 8\} ... \{X_n ... 8\}$. Therefore, under (H1), we know that $X\tilde{}$ is positive recurrent with stationary distribution $\mu\tilde{}$, so that in particular $P(X_n ... 8) = 0$. Furthermore, observe that for all $t > 0$, $u = 0$:

$$\int_{s=0}^{t} I_{\{X(s)=u\}} ds = \sum_{k=1}^{N(t)} \int_{s=T_{k-1}}^{T_k} I_{\{X(s)=u\}} ds + \int_{s=T_{N(t)}}^{t} I_{\{X(s)=u\}} ds.$$

Therefore, for all $k ... N$, $\int_{s=T_k}^{T_{k+1}} I_{\{X(s)=u\}} ds = I_{\{X_k=u\}} \cdot ...T_{k+1} ... \log(X_k/u)\cdot\cdot k$.

Now, applying the strong law of large number (SLLN) to the positive recurrent chain $((X_n, \cdot\cdot n, ...T_{n+1}))_{n...N}$ with invariant probability distribution $\mu\tilde{}(dx)...H(d..)...G(dt)$, we get that

$$n^{-1} \sum_{k=1}^{n} \int_{s=T_k}^{T_{k+1}} I_{\{X(s)=u\}} ds ... \int_{x=u} \int_{.....T} \int_{t=0}^{8} \int_{..}^{8} \log(x/u)\, t \, \mu\tilde{}(dx)H(d..)G(dt). \tag{21}$$

As we assumed $m_G = E(...T_k) < 8$ for $k = 2$, we have the following convergence for the delayed renewal process: $N(t)/t ... m_G^{-1}$ as $t ... 8$. Combined with (21), this yields $t^{-1} \int_{s=0}^{t} I_{\{X(s)=u\}} ds ... \mu([u, 8[)$ as $t ... 8$, with μ given by (11).

We thus proved that X(t) has a limiting probability distribution μ, which has density f (y) given by

$$f(y) = m_G^{-1} \int_{.....T}^{Z} \int_{t=0}^{8} \tilde{f}(ye^{..t}) e^{..t} \bar{G}(t) dt H(d..),$$

denoting by $\bar{G} = 1 - G$ the inter−intake survival function.

Besides, if sup T < 8, from (11), we immediately have that, for all u > 0, t > 0,

$$m_G^{t \,...\, \log 2 / \sup T} \bar{G}(t) \mu\tilde{}([2u, 8[) = \mu([u, 8[) = \mu\tilde{}([u, 8[).$$

The distributions μ and μ~ have thus exactly the same right tail behavior.

Assuming now that (H2) and (H4) are both fulfilled, we turn to the study of the trivariate process $\{(X(t), ..(t), A(t))\}_{t=0}$. It may be easily seen as .. ... H ... G− irreducible and any compact set $[0, s]\times[0, \bar{..}]\times[0, a]$, with $s > 0$, $\bar{..} > 0$, $a > 0$ is a 'pe− tite set' for the latter (see [43] for an account of stochastic stability concepts related to continuous−time Markov processes). Indeed, denote by $Q_t(. \mid (x_0, ..0, a_0))$ the distribution of (X(t), ..(t), A(t)) conditioned upon (X(0), ..(0), A(0)) = $(x_0, ..0, a_0)$. One may easily check that its trace on the event {N (t) = 1} has density $f_U(e^{..a} x - x_0 e^{-.. 0(t-a)}) e^{..a} \bar{G}(a) g_{a_0}(t - a)$ with respect to ..(dx) ... H(d..) ... ..(da). Hence, for all $(x_0, ..0, a_0)$ ... [0, s] × [0, $\bar{..}$] × [0, a], we have the minorization condition:

$$Q_t(. \mid (x_0, ..0, a_0)) = d(s, \bar{..}, a) \cdot U_s \,...\, H(. \cap [0, \bar{..}]) \,...\, U_a, \tag{22}$$

with $d(s, \bar{..}, a) = (s \times H([0, \bar{..}]) \times a) \cdot \inf_{x...[0, se^{..a}]} f_U(x) \cdot \inf_{v...[t-a,t]} g(v) \cdot \bar{G}(a)$.

Following [38] (see Theorem 4.1 therein), let .., d be such that $0 < .. < d$ and consider the Lyapounov function $V(x, .., t) = (1 + x\,..)(1 + ..)W(t)$ on R^{3+} with $W(t) = 1 + G(t)e^{-..t}\left[1 + \int_{x=t}^{8} e^{dx} \frac{G(x)}{G(t)} dx\right]$ (notice that, under (H4), W (t) ... 8 as t ... 8). It may be easily seen that the test function V belongs to the domain of the infinitesimal generator G (see Eq. (5)) and that $GV(x, .., t) = -..V(x, .., t) + b(x, .., t)$, where

$$\begin{aligned} b(x, .., t) &= (1 + x\,..)(1 + ..)[z(t) - ....x\,..W(t)] \\ &\quad - ..(t)[(1 + x\,..)(1 + ..) - (1 + E[..])(1 + E[(x + U)..])], \\ z(t) &= .. + \frac{g(t)}{\bar{G}(t)} e^{-..t}\left[1 - 2G(t) + \int_{u=t}^{Z} e^{du} \bar{G}(u)du - G(t)e^{(d-..)t}\right]. \end{aligned}$$

Observe that b(x, .., t) ... −8 as (x, .., t) tends to infinity. Hence, there exist $s > 0$, $\bar{..} > 0$ and $a > 0$ large enough and $b < 8$ such that the following drift condition holds:

$$GV(x, .., t) = -..V(x, .., t) + bI_{\{(x,..,t)...[0,s]\times[0,\bar{..}]\times[0,a]\}}. \tag{23}$$

Then, (12) directly follows from Theorem 5.3 in [41].

A.3. Proof of Proposition 1. 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 on the right−hand side of (24) being bounded by $x_0 T_1/T$, it almost surely converges to 0 as $T \to \infty$. Also, we have for all $k = 1$,

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

Furthermore, by virtue of Theorem 3.1, assumption (H2) with $\beta = 1$ ensures that $m_{\tilde{\mu}} = \int_{x=0}^{\infty} x\tilde{\mu}(dx) < \infty$ and consequently that

$$m_{\tilde{\mu}} = \int_{x=0}^{\infty}\int_{t=0}^{\infty}\int_{\alpha=T_1}^{\infty} \frac{x(1 - e^{-\alpha t})}{\alpha}\tilde{\mu}(dx)H(d\alpha)G(dt) < \infty,$$

making the SLLN for the positive recurrent chain $((X_n, \alpha_n, T_{n+1}))_{n\geq 1}$ applicable to $\sum_{n=1}^{\infty}(1 - \exp(-\alpha_n (T_{n+1}-T_n)))X_n/\alpha_n$ (refer to Theorem 17.3.2 in [44] for instance). We thus have that

$$N^{-1}\sum_{k=1}^{N}\frac{X_k}{\alpha_k}(1 - e^{-\alpha_k (T_{k+1}-T_k)}) \to m_{\tilde{\mu}}\int_{t=0}^{\infty}\int_{\alpha=T_1}^{\infty}\frac{1 - e^{-\alpha t}}{\alpha}H(d\alpha)G(dt) \quad \text{a.s.}, \tag{25}$$

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 (13). Notice that $m_\mu = \int_{t=0}^{\infty}\int_{\alpha=T_1}^{\infty}(1 - \exp(-\alpha t))/\alpha H(d\alpha)G(dt) m_{\tilde{\mu}}/m_G$.

We now turn to the proof of the Central Limit Theorem (CLT). Using again Theorem 3.1, we have that $\int x^2\tilde{\mu}(dx) < \infty$ when (H2) holds for some $\beta = 2$, so that

$$\int_{x=0}^{\infty}\int_{t=0}^{\infty}\int_{\alpha=T_1}^{\infty} \frac{x^2(1 - e^{-\alpha t})^2}{\alpha^2}\tilde{\mu}(dx)H(d\alpha)G(dt) < \infty.$$

By virtue of the CLT for positive recurrent chains (see Theorem 17.0.1 in [44]), we have that $N^{-1/2}\sum_{k=1}^{N}\{(1 - e^{-\alpha_k (T_{k+1}-T_k)})X_k/\alpha_k - m_{\tilde{\mu}}\}$ converges in distribution to $\mathcal{N}(0, \sigma_{\tilde{\mu}}^2)$ as $N \to \infty$, with $\sigma_{\tilde{\mu}}^2 = E_\mu[(X_1(1-e^{-\alpha_1 (T_2-T_1)})/\alpha_1 - m_{\tilde{\mu}})^2] + 2\sum_{k=2}^{\infty} E_\mu[(X_1(1-e^{-\alpha_1 (T_2-T_1)})/\alpha_1 - m_{\tilde{\mu}})(X_k(1-e^{-\alpha_k (T_{k+1}-T_k)})/\alpha_k - m_{\tilde{\mu}})]$.

One may then easily deduce (14) from (24) with $\sigma^2 = \sigma_{\tilde{\mu}}^2/m_G$.

A.4. Proof of Theorem 4.1. Observe first that (16) immediately follows from (15) by virtue of standard properties of Wasserstein metrics. In order to prove (15), we construct a specific coupling of the laws $\mathcal{L}_{X^{(T)}}$ and $\mathcal{L}_{\hat{X}^{(T)}}$. Let $(V_k)_{n\geq N}$, $(V_k')_{k\geq N}$ and $(V_k'')_{k\geq N}$ 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

$$\Delta T_k = G^{-1}(V_k), \quad U_k = F_U^{-1}(V_k'), \quad \alpha_k = H^{-1}(V_k''),$$

$$\Delta T_k^{(n)} = G^{(n)-1}(V_k), \quad U_k^{(n)} = F_U^{(n)-1}(V_k'), \quad \alpha_k^{(n)} = H^{(n)-1}(V_k''),$$

and define recursively for $k \geq N$, $X_{k+1} = X_k e^{-\alpha_k \Delta T_{k+1}} + U_{k+1}$ and $X_{k+1}^{(n)} = X_k^{(n)} e^{-\alpha_k^{(n)} \Delta T_{k+1}^{(n)}} + U_{k+1}^{(n)}$ with $X_0 = X_0^{(n)} = x_0$, as well as $T_{k+1} = \Delta T_{k+1} + T_k$ and $T_{k+1}^{(n)} = \Delta T_{k+1}^{(n)} + T_k^{(n)}$ with $T_0 = T_0^{(n)} = 0$. For notational convenience, the su−perscript (n) is omitted in the sequel. Using in particular the fact that $x - 0.7 \leq e^{-x}$

is 1−Lipschitz, straightforward computations yield

$$\hat{X}_k - X_k = x_0\Big\{\sum_{i=1}^{k} \cdot\cdot_i \ldots \hat{T}_{i+1} - \ldots \hat{T}_{i+1} + \sum_{i=1}^{k} \ldots \hat{T}_{i+1} \cdot\cdot_i - \cdot\cdot\hat{}_i\Big\}$$
$$+ \sum_{i=1}^{k} U_i \Big(\sum_{j=i}^{k-1} \cdot\cdot_j \ldots \hat{T}_{j+1} - \ldots \hat{T}_{j+1} + \sum_{j=i}^{k-1} \ldots \hat{T}_{j+1} \cdot\cdot_j - \cdot\cdot\hat{}_j\Big) + \sum_{i=1}^{k} \hat{U}_i - U_i \tag{26}$$

Turning now to the coupling construction in continuous time, define $N(t) = \sum_{k=1} I_{\{T_k=t\}}$ and $\hat{N}(t) = \sum_{k=1} I_{\{\hat{T}_k=t\}}$, as well as $X(t) = X_{N(t)} \exp(-\cdot\cdot_{N(t)}(t - T_{N(t)}))$ and $\hat{X}(t) = \hat{X}_{\hat{N}(t)} \exp(-\cdot\cdot\hat{}_{\hat{N}(t)}(t - \hat{T}_{\hat{N}(t)}))$ for $t = 0$. Set also $T_k^+ = T_k \ldots \hat{T}_k$ and $T_k^- = T_k \ldots \hat{T}_k$ for all $k \ldots N$ and observe that

$$m_H(G_{\hat{X}(T)}, G_{X(T)}) = \max_{0=k=N(T)\ldots\hat{N}(T)} M_k, \tag{27}$$

where

$$M_k = \sup_{T_k^+=t<T_{k+1}^-} X(t - (\hat{T}_k - T_k)_+) - \hat{X}(t - (T_k - \hat{T}_k)_+) + T_k - \hat{T}_k$$
$$+ \sup_{T_{k+1}^-=t<T_{k+1}^+} X(t - (t - T_{k+1})_+) - \hat{X}(t - (t - \hat{T}_{k+1})_+) + T_{k+1} - \hat{T}_{k+1},$$

denoting by $x_+ = 0 \ldots x$ the positive part of any $x \ldots R$. It follows from easy calculations that

$$M_k = X_k - \hat{X}_k + X_k(\ldots T_{k+1} \ldots \ldots \hat{T}_{k+1}) \cdot\cdot_k - \cdot\cdot\hat{}_k + \sum_{i=1}^{k} \ldots T_i - \ldots \hat{T}_i$$
$$+ X_{k+1} - \hat{X}_{k+1} + U_{k+1} - \hat{U}_{k+1} + \sum_{i=1}^{k+1} \ldots T_i - \ldots \hat{T}_i.$$

By taking the expectation in (27) and then using the bounds $X_k = x_0 + \sum_{1=i=k} U_i$ and (26) combined with Wald's lemma, straightforward computations yield

$$E[m_{M_2^{(T)}}(\hat{X}(T), X(T))] = (1 + E[N(T) \ldots \hat{N}(T)])\{2x_0(m_H M_1(G, \hat{G}) + m_{\hat{G}} M_1(H, \hat{H})$$
$$+ 3M_1(F_U, \hat{F}_U) + 2M_1(G, \hat{G}) + (x_0 + m_{F_U})(T + m_G + m_{\hat{G}})$$
$$\times M_1(H, \hat{H})\} + E[(1 + N(T) \ldots \hat{N}(T))^2] m_{F_U}(m_H M_1(G, \hat{G})$$
$$+ m_{\hat{G}} M_1(H, \hat{H})),$$

denoting by m_F (resp. $m_{\hat{F}}$) the mean of the distribution function F (resp. of the estimate $\hat{F}$), F being any of the distribution functions G, F_U or H (notice that $m_{\hat{F}} = v_F + m_F$). In addition, there exist constants C, $C' < 8$ s.t. $E(N(T)) \ldots E(\hat{N}(T)) = CT$ and $E(N(T)^2) \ldots E(\hat{N}(T)^2) = C'T^2$ (refer to Propositions 6.1 and 6.3 of chap. V in [4] for instance). Observe that the constants C and C' may be chosen independently from the integer n indexing the sequence $\hat{G}$, since by assumption the sequences $m_{\hat{G}}$ and $s^2_{\hat{G}}$ are bounded. This establishes the desired result (15).

Acknowledgements. We would like to thank the referee and the editor for their valuable comments and suggestions.

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 (1969), no. 4, 478–484.
2. L. S. Allen, An introduction to stochastic processes with biology applications, Pearson/Prentice Hall, Upper Saddle River, N.J., 2003.
3. S. Asmussen, Subexponential asymptotics for stochastic processes: Extremal behavior, station−ary distributions and first passage probabilities, Adv. Appl. Probab. 8 (1998), no. 2, 354–374.
4. , Applied probability and queues, Springer−Verlag, New York, 2003, Second Edition.
5. P. Bertail and S. Cle´menc¸on, Edgeworth expansions for suitably normalized sample mean statistics of atomic Markov chains, Prob. Th. Rel. Fields 130 (2004), no. 3, 388–414.
6. , Regenerative−block bootstrap for Markov chains, Bernoulli 12 (2005), no. 4, 689–712.
7. P. Bertail, S. Cle´menc¸on, and J. Tressou, Statistical analysis of a dynamic model for food contaminant exposure with applications to dietary methylmercury contamination, 2007, http: //metarisk.inapg.inra.fr/.
8. P. Bertail, M. Feinberg, J. Tressou, and P. Verger (Coordinateurs), Analyse des risques ali−mentaires, TEC&DOC, Lavoisier, Paris, 2006.
9. P. Bertail and J. Tressou, Incomplete generalized U−Statistics for food risk assessment, Bio−metrics 62 (2006), no. 1, 66–74.
10. P. J. Bickel and D. A. Freedman, Some asymptotic theory for the bootstrap, Ann. Statist. 9 (1981), 1196–1217.
11. P. Bougerol and N. Picard, Strict stationarity of generalized autoregressive processes, Ann. Probab. 20 (1992), no. 4, 1714–1730.
12. M. Brett, H .J. Weimann, W. Cawello, H. Zimmermann, G. Pabst, B. Sierakowski, R. Giesche, and A. Baumann, Parameters for compartment−free pharmacokinetics: Standardisation of study design, data analysis and reporting, Shaker Verlag, Aachen, Germany, 1999.
13. P. J. Brockwell, S. I. Resnick, and R. L. Tweedie, Storage processes with general release rule and additive inputs, Adv. Appl. Probab. 14 (1982), no. 2, 392–433.
14. S. Browne and K. Sigman, Work−modulated queues with applications to storage processes, J. Appl. Probab. 29 (1992), no. 3, 699–712.
15. E. Budtz−Jrgensen, N. Keiding, P. Grandjean, P. Weihe, and R. F. White, Statistical methods for the evaluation of health effects of prenatal mercury exposure, Environmetrics 13 (2003), 105–120.
16. F. Ce´rou, P. Del Moral, F. Le Gland, and P. Lezaud, Genetic genealogical models in rare event analysis, Alea 1 (2006), 183–196.
17. O. L. V. Costa, Stationary distributions for piecewise−deterministic Markov processes, J. Appl. Probab. 27 (1990), no. 1, 60–73.
18. 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 (1989), 318–332.
19. CREDOC−AFSSA−DGAL, Enque^te INCA (individuelle et nationale sur les consommations alimentaires), TEC&DOC, Lavoisier, Paris, 1999, (Coordinateur : J.L. Volatier).
20. A. Cre´pet, J. Tressou, P. Verger, and J. Ch. Leblanc, Management options to reduce expo−sure to methyl mercury through the consumption of fish and fishery products by the French population, Regul. Toxicol. Pharmacol. 42 (2005), 179–189.
21. D. J. Daley and D. Vere−Jones, An introduction to the theory of Point Processes: I, Springer−Verlag, New York, 2003, Second Edition.
22. M. H. A. Davis, Piecewise−deterministic Markov processes: A general class of non−diffusion stochastic models, J. R. Statist. Soc. 46 (1984), no. 3, 353–388.
23. M. H. A. Davis and R. J. Elliott, Applied stochastic analysis, Stochastics Monographs, vol. 5, Gordon and Breach, New York, 1991.
24. L. Edler, K. Poirier, M. Dourson, J. Kleiner, B. Mileson, H. Nordmann, A. Renwick, W. Slob, K. Walton, and G. Wu¨rtzen, Mathematical modelling and quantitative methods, Food Chem. Toxicol. 40 (2002), 283–326.
25. P. Embrechts, C. Klu¨ppelberg, and T. Mikosch, Modelling extremal events for insurance and finance, Applications of Mathematics, Springer−Verlag, Berlin, 1997.
26. 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

27. FAO/WHO, Evaluation of certain food additives and contaminants, Fifty−seventh report of the Joint FAO/WHO Expert Committee on Food Additives 909, 2002, June 2001.
28. , Evaluation of certain food additives and contaminants, Sixty−first report of the Joint FAO/WHO Expert Committee on Food Additives, Technical Report Series 52, WHO, Geneva, Switzerland, 2004, June 2003.
29. M. Gibaldi and D. Perrier, Pharmacokinetics, Drugs and the Pharmaceutical Sciences: a Series of Textbooks and Monographs, Marcel Dekker, New York, 1982, Second Edition.
30. 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 (2003), no. Suppl. 1, S1–S7.
31. 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 (1997), 417–428.
32. D. R. Grey, Regular variation in the tail behaviour of solutions of random difference equations, Adv. Appl. Probab. 4 (1994), no. 1, 169–183.
33. D. R. Helsel, Nondetects and data analysis : Statistics for censored environmental data, Statistics in Practice, Wiley, New York, 2004.
34. IFREMER, Re´sultat du re´seau national d'observation de la qualite´ du milieu marin pour les mollusques (RNO), 1994−1998.
35. IPCS, Environmental health criteria 70: Principles for the safety assessment of food, Tech. report, WHO, Geneva, 1987, International Programme on Chemical Safety, available at www.who.int/pcs/.
36. J. Jacod and A. N. Shiryaev, Limit theorems for stochastic processes, Springer, New York, 1987.
37. H. Kesten, Random difference equations and renewal theory for products of random matrices, Acta Math. 131 (1973), 207–248.
38. T. Konstantopoulos and G. Last, On the use of Lyapunov function methods in renewal theory, Stoch. Proc. App. 79 (1999), 165–178.
39. R. Lund, S. P. Meyn, and R. L. Tweedie, Rates of convergence of stochastically monotone Markov processes, Ann. App. Prob. 6 (1996), 218–237.
40. MAAPAR, Re´sultats des plans de surveillance pour les produits de la mer, Ministe`re de l'Agriculture, de l'Alimentation, de la Pe^che et des Affaires Rurales, 1998−2002.
41. S. P. Meyn and R. L. Tweedie, Stability of Markovian processes III: Foster−Lyapunov criteria for continuous time processes, Adv. Appl. Probab. 25 (1993), no. 3, 518–548.
42. , Computable bounds for geometric convergence rates of Markov chains, Ann. Appl. Probab. 4 (1994), no. 4, 981–1011.
43. , A survey of Foster−Lyapunov conditions for general state space Markov processes, Proceedings of the Workshop on Stochastic Stability and Stochastic Stabilization (Metz, France), Springer−Verlag, 1994.
44. , Markov chains and stochastic stability, Springer−Verlag, London, 1996.
45. J. K. Miettienen, Mercury, mercurials, and mercaptans, M.W. Miller & T.W. Clarkson ed., ch. Absorption and elimination of dietary mercury and methyl mercury in man, pp. 223–240, C.C. Thomas, Springfield, IL., 1973.
46. S. T. Rachev, Probability metrics and the stability of stochastic models, Wiley, New York, 1991.
47. S. T. Rachev and G. Ru¨schendorf, Mass transportation problems I: Theory, Springer, New York, 1998.
48. S. T. Rachev and G. Samorodnitsky, Limit laws for a stochastic process and random recursion arising in probabilistic modelling, Adv. Appl. Probab. 27 (1995), no. 1, 185–202.
49. T. Rahola, R. K. Aaron, and J. K. Miettienen, Half time studies of mercury and cadmium by whole body counting, Assessment of Radioactive Contamination in Man, Proceedings IAEA−SM−150−13, 1972, pp. 553–562.
50. A. G. Renwick, S. M. Barlow, I. Hertz−Picciotto, A. R. Boobis, E. Dybing, L. Edler, G. Eisen−brand, J. B. Greig, J. Kleiner, J. Lambe, D. J. G. Mueller, S. Tuijtelaars, P. A. Van den Brandt, R. Walker, and R. Kroes, Risk characterisation of chemicals in food and diet, Food Chem. Toxicol. 41 (2003), no. 9, 1211–1271.
51. G. O. Roberts and R. L. Tweedie, Bounds on regeneration times and convergence rates of Markov chains, Stoch. Proc. Appl. 80 (1999), 211–229.

52. , Rates of convergence of stochastically monotone and continuous time Markov models, J. Appl. Probab. 37 (2000), no. 2, 359–373.
53. J. S. Rosenthal, Minorization conditions and convergence rates for Markov chain Monte Carlo, J. Amer. Stat. Assoc. 90 (1995), 558–566.
54. M. Rowland and T. N. Tozer, Clinical pharmacokinetics: Concepts and applications, 1995.
55. 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 (1994), 251–256.
56. J. C. Smith and F. F. Farris, Methyl mercury pharmacokinetics in man: A reevaluation, Toxicol. Appl. Pharmacol. 137 (1996), 245–252.
57. J. Tressou, Me´thodes statistiques pour l'e´valuation du risque alimentaire, Ph.D. thesis, Universite´ Paris X, 2005, http://tel.archives−ouvertes.fr/tel−00139909.
58. , Non parametric modelling of the left censorship of analytical data in food risk exposure assessment, J. Amer. Stat. Assoc. 101 (2006), no. 476, 1377–1386.
59. J. Tressou, A. Cre´pet, 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 (2004), no. 8, 1349–1358.
60. T. Tsubaki and K. Irukayama, Minamata disease: Methyl−mercury poisoning in Minamata and Niigata, Japan., ElsevierScientific, New York, 1977.
61. NRC (National Research Council), Committee on the toxicological effects of methylmercury: Toxicological effects of methylmercury, Tech. report, Washington DC, 2000.
62. P. Verger, J. Tressou, and S. Cle´menc¸on, Integration of time as a description parameter in risk characterisation: application to methylmercury, Regul. Toxicol. Pharmacol. 49 (2007), no. 1, 25–30.
63. W. Whitt, Stochastic−process limits. an introduction to stochastic−process limits and their application to queues, Springer, New York, 2002.

Received on April 10, 2007. Accepted on October 12, 2007.

E−mail address: Stephan.Clemencon@agroparistech.fr
E−mail address: Patrice.Bertail@ensae.fr Jessica.Tressou@agroparistech.fr