Lévy processes and stochastic von Bertalanffy models of growth, with application to fish population analysis

Tommaso Russo [a,*], Paolo Baldi [b], Antonio Parisi [c], Giuseppe Magnifico [d], Stefano Mariani [e], Stefano Cataudella [a]

[a] *Laboratorio di Ecologia Sperimentale e Acquacoltura, Dipartimento di Biologia, Università di Roma Tor Vergata, Italy*
[b] *Dipartimento di Matematica, Università di Roma Tor Vergata, Italy*
[c] *Dipartimento di Studi Economico-Finanziari e Metodi Quantitativi Facoltá di Economia, Università di Roma Tor Vergata, Italy*
[d] *Consiglio Nazionale delle Ricerche, CNR, Roma, Italy*
[e] *School of Biology and Environmental Science, University College of Dublin, Ireland*

ARTICLE INFO

ABSTRACT

The study of animal growth is a longstanding crucial topic of theoretical biology. In this paper we introduce a new class of stochastic growth models that enjoy two crucial properties: the growth path of an individual is monotonically increasing and the mean length at time t follows the classic von Bertalanffy model. Besides the theoretical development, the models are also tested against a large set of length-at-age data collected on Atlantic herring (*Clupea harengus*): the mean lengths and variances of the cohorts were directly estimated by least squares. The results show that the use of subordinators can lead to models enjoying interesting properties, in particular able to catch some specific features often observed in fish growth data. The use of subordinators seems to allow for an increased fidelity in the description of fish growth, whilst still conforming to the general parameters of the traditional von Bertalanffy equation.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

The modelling of growth and the analysis of intra-population pattern of size variability through time are the central topics in animal population biology, since the internal size structure of populations can have a decisive influence on the population dynamics (DeAngelis et al., 1993; Imsland et al., 1998; Uchmanski, 2000; Kendall and Fox, 2002; Fujiwara et al., 2004). In general, the von Bertalanffy growth function (VBGF, von Bertalanffy, 1957) is the best acknowledged and used relationship to describe the growth of fish and other animals. This equation states that the size of an individual increases in time according to the equation

$$x_t = L_\infty \left(1 - e^{-k(t+t_0)}\right), \tag{1.1}$$

where L_∞ is the extremal length that is attained as time goes to infinity, $-t_0$ is the time of conception, at which the size should be 0 and k is a parameter that gives the speed of the process: the larger the value of k, the quickest the growth. The VBGF is most commonly used as a descriptive model of size-at-age data (Essington et al., 2001). Nevertheless, Eq. (1.1) describes the relationship between age and mean length of a population,

whereas the variability among individuals of the same age (e.g. the variance or even the distribution of each cohort) is not included.

The popular assumption of Gaussianity (Imsland et al., 1998) is clearly a first (rough) approximation in this direction. A natural approach to the problem of determining the appropriate form of the probability distribution for a population at a given time t is to model first the growth process of the individuals (individual-based models, IBM). Nowadays, in both ecological (Arino et al., 2004), evolutionary (Conover and Munch, 2002; Ernande et al., 2004) and management (Caswell, 2001) contexts, one of the challenges of the researcher is to model how the size of an individual changes in time and to deduce from the growth model which kind of probability distribution models the size of fish at a given age (Lv and Pitchford, 2007; Fujiwara et al., 2004).

A suitable model of growth should account for both individual and environmental variability. In fish, as in other animals, the first source of variability is rooted in the physiological processes and is the net result of two opposing processes, catabolism and anabolism (von Bertalanffy, 1938). The inter-individual variability in growth is the result of several internal (genetic) and external (environmental) factors which affect these physiological processes. In fact, whilst each individual is born with a personal genetic architecture which primarily determines his growth profile, a number of physical and biological factors, such as water

* Corresponding author. Tel.: +390672595974; fax: +390672595965.
E-mail address: Tommaso.Russo@Uniroma2.it (T. Russo).

temperature (Sumpter, 1992), dissolved oxygen (Brett, 1979), photoperiod (Imsland et al., 2002), and the availability of appropriate food sources (Rilling and Houde, 1999), have been shown to affect growth rates. In order to take into account these aspects as well as individual variability, a class of IBM was developed (Lv and Pitchford, 2007; Sainsbury, 1980; Mulligan and Leaman, 1992; Wang and Thomas, 2003; Imsland et al., 2002; Wang, 1999; Gudmundsson, 2005).

These models can be classified into two main categories. The first comprises those which consider the inter-individual variability as a stochastic factor to be added to the general growth curve of the population. The distribution of this factor is the same for all the individuals. In the most recent approaches of this kind (Gudmundsson, 2005; Lv and Pitchford, 2007; Wang, 1999) some individual-based stochastic models of growth are proposed using a stochastic differential equations. These models take the general form

$$dL_t = f(L_t, t) + \alpha(L_t, t) \, dB(t). \tag{1.2}$$

Here L_t is the size at time t, $f(L_t, t)$ characterizes the deterministic intrinsic growth (drift coefficient) of the individual (the same for all individuals); $\alpha(L_t, t)$ gives the magnitude of the random fluctuations (diffusion coefficient) and $B(t)$ is a standard Brownian motion, or Wiener process, which is commonly used to model a variety of background and environmental fluctuations in physical, financial and biological contexts (see, e.g. Karlin and Taylor, 1981).

The stochastic component of these models is intended to account for both the environment and the inter-individual variability. It should be stressed, however, that (as already remarked in Gudmundsson, 2005) the solution of an equation as Eq. (1.2) cannot be monotonically increasing and therefore appears to be unsuitable to model the evolution of the size of an individual. Stochastic models like the one of Eq. (1.2) are conversely well suited in order to model quantities, as prices in financial markets, that are characterized by an oscillating and therefore non-monotonic behaviour. This aspect seems to be a drawback for a growth model of several animals like fish. In fact, for fish and other vertebrates, the physiological mechanism of growth in length (i.e. the addition of bone material to the axil unit of the skeleton that is the vertebra) leads to a pattern in which the size of an individual is necessarily increasing in time (Weatherley and Gill, 1987). Let us point out that there exist no Gaussian process which is increasing. Therefore, modelling size variability through an individual growth process cannot give rise to a Gaussian distribution.

Closely related to the models (Eq. (1.2)) are those introduced in Gudmundsson (2005) where it is the derivative of the growth process that is the solution of a stochastic equation. It is possible in this way to obtain a stochastic process that is increasing. These models are interesting and deserve to be tried by testing against real data. Remark, however, that the mean size at time t of a population following such a model does not follow a VBGF.

The second category of stochastic models suggested so far comprises non-deterministic models in which the individuals of a fish population have different parameters of the VBGF. In this way, each individual has its own triplet (L_∞, t_0, k), which is retained throughout its life (Sainsbury, 1980). Considering the length to age relationship, the length X_t at age t of the ith individual with the parameters triple $(L_i, t_{0,i}, k_i)$ is given by

$$X_{t,i} = L_i \left(1 - e^{-k_i(t + t_{0,i})} \right). \tag{1.3}$$

This model displays a monotonic behaviour and considers the inter-individual variability of growth parameters, but does not account for the randomness coming from the environment, which can be seen as a limitation of the model.

Finally, it should be considered that individual growth is a complex energetic process. Individual length increases only when enough energy from food is available for growth. On the other hand the lack of food does not usually lead to a reduction in length, because organisms can lose body mass without shrinking in length (Kooijman, 2000) which also because of the presence of a skeleton. Energy may also be allocated to storage for future use, producing "memory" in growth dynamics. Individual organisms encounter and ingest food, which is then assimilated. Assimilated food is transformed into reserve material such as protein and fat. A fixed fraction of the energy from the reserve is used for both metabolic maintenance and growth, and the rest is used for reproduction. Looking at the length of an individual, we suggest that this process determines a pattern in which periods of no growth (determined by scant energy inlet) are separated by periods of growth. If the periods of growth are short, the growth process could be well described by a model allowing for discontinuities, i.e. for instantaneous increases of the length (jumps). This idea is consistent with several observations reported for fish in general and for the species we are going to study (Hinrichsen et al., 2007).

In this paper we introduce two classes of stochastic models of growth that attempt to overcome the drawbacks pointed out above. The main idea developed in this paper is to model the growth process as the solution of a stochastic equation of the form

$$dX_t = (L_\infty - X_{t-}) \, dZ_t,$$

where $(Z_t)_t$ is a *subordinator*. These are a class of stochastic processes that are strictly increasing and the solution $(X_t)_t$ turns out to be increasing also. These models enjoy a certain number of desirable features, namely

- they take into account both the individual and the environmental sources of randomness;
- they are increasing;
- the mean size at time t follows a VBGF.

In Section 2 we make a quick review on the topics of subordinators upon which our models are built. In Sections 3 and 4 some models are developed, attempting to answer to the points exposed above. Finally in Section 6 we apply the proposed models to a large set of length-at-age data of Atlantic herring *Clupea harengus*, presented in Section 5.

2. Subordinators

A *subordinator* is a stochastic process $(Z_t)_t$ such that

- $Z_0 = 0$;
- its paths are right continuous and increasing almost surely;
- has independent and stationary increments.

This means in particular that the distribution of $Z_{t+h} - Z_t$ is independent of t for every $h \geqslant 0$ and that $Z_t - Z_s$ and $Z_v - Z_u$ are independent r.v.'s for $u < v \leqslant s < t$. Also the increments $Z_{t+h} - Z_t$ must be stationary, in the sense that their distribution depends on h only and not on t. The characterization of such processes (that are particular instances of Lévy processes) has received much attention in time and it is characterized in terms of the Laplace transform of Z_t. It is immediate that if

$$M_t(\theta) = E(e^{-\theta Z_t}), \quad \theta \geqslant 0 \tag{2.1}$$

then M is of the form

$$M_t(\theta) = e^{t\phi(\theta)}, \tag{2.2}$$

where the exponent ϕ is characterized by the Lévy–Khintchin formula (see for the subjects developed in this section and in the next ones, e.g., Cont and Tankov, 2004; Sato, 1999). More precisely, the Lévy–Khintchin formula for subordinators states that

$$\phi(\theta) = -\theta\gamma + \int_0^{+\infty} (e^{-\theta x} - 1) v\,(dx),$$

where $\gamma \geqslant 0$ and v, the *Lévy measure*, is a positive measure on $\mathbb{R}^+$ such that

$$\int_0^{+\infty} \frac{x}{1+x} v\,(dx) < +\infty. \tag{2.3}$$

Intuitively a subordinator increases as the superposition of a deterministic evolution $t \to \gamma t$ and of a stochastic process which only makes jumps. These are made at times governed by a Poisson process. More precisely, if $0 < a < b$, then $v([a,b])$ is the intensity of the Poisson process of the jumps whose size is larger than a and smaller than b. This means that if t is the (random) time at which the jump occurs, then the path has, at t, a left limit X_{t-} and a right limit X_{t+} that are different, with $X_{t-} < X_{t+}$. For our purposes we shall therefore be much interested in subordinators whose Lévy measure is large to near 0 and gives a small mass only away from 0. The resulting process would hence be characterized by many very small jumps.

As every Lévy process, the independence and stationarity properties of the increments imply that if Z_t has a finite mean, then

$$E(Z_t) = \mu t$$

for some real number μ and if it has a finite variance then

$$\mathrm{Var}(Z_t) = \sigma^2 t$$

for some $\sigma^2 \geqslant 0$.

It is not possible for a general Lévy measure to compute explicitly the Laplace transform and/or the density at time t of the associated subordinator. The next examples introduce some families of subordinators of interest for which these features can be determined.

- The Gamma process, where the distribution of Z_t follows a $\Gamma(\alpha t, \lambda)$ distribution. Recall that such a distribution has a density

$$f(x) = \frac{\lambda^{\alpha t}}{\Gamma(\alpha t)} x^{\alpha t - 1} e^{-\lambda x}, \quad x > 0$$

and $f(x) = 0$ for $x \leqslant 0$. Of course the increment $Z_t - Z_s$, having the same distribution as Z_{t-s}, has a $\Gamma(\alpha(t-s), \lambda)$ distribution. Its Laplace transform is

$$M_t(\theta) = \left(\frac{\lambda}{\lambda + \theta}\right)^{\alpha t} \tag{2.4}$$

so that $\phi(\theta) = -\alpha \log \lambda / \lambda + \theta$. From the well-known properties of the $\Gamma(\alpha t, \lambda)$ distribution the statistical indices of interest of a Gamma process are: mean $= \alpha t / \lambda$, variance $= \alpha t / \lambda^2$, skewness $= 2/\sqrt{\alpha t}$ and kurtosis $= 3 + 6/(\alpha t)$.
Here the Lévy measure is

$$v\,(dy) = \frac{\alpha}{y} e^{-\lambda y}\,dy.$$

This process therefore gives a high intensity to small jumps (see Fig. 1).
- The *inverse Gaussian processes*, where Z_t is defined as the first time at which a process of the form $\sqrt{\rho} B_s + bs$ crosses level t. Here $(B_t)_t$ denotes a standard Brownian motion and $\rho, b > 0$. Its
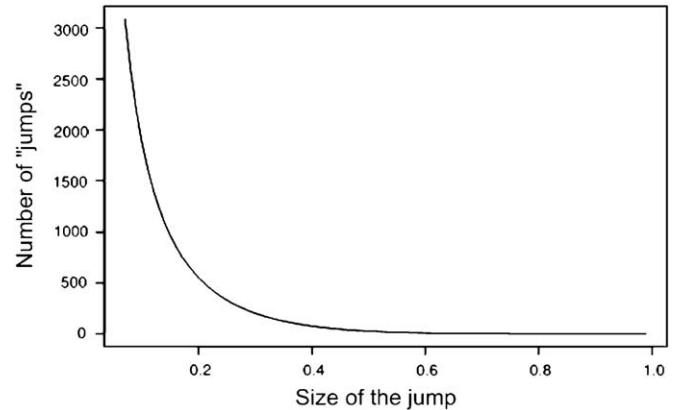

Fig. 1. The graph of the Lévy density h of Eq. (4.7) for $\alpha = 300$, $\lambda = 1.2$. The intensity of small jumps diverges at 0 and h vanishes at 1, so that large jumps are unlikely or forbidden.

density at time t is, for $x > 0$,

$$f_t(x) = \frac{t\,e^{bt/\rho}}{(2\pi\rho)^{1/2} x^{3/2}} \exp\left(-\frac{b^2}{2\rho}x - \frac{t^2}{2\rho x}\right).$$

It depends on the positive parameters b, ρ. Its Laplace transform at time t is

$$M_t(\theta) = \exp\left(-t\left(\sqrt{\frac{b^2}{\rho^2} + \frac{2\theta}{\rho}} - \frac{b}{\rho}\right)\right),$$

from which one can derive the expression of the statistical indices of interest: mean $= t/b$, variance $= \rho t / b^3$, skewness $= 3\sqrt{\rho/(tb)}$ and kurtosis $= 3 + 15\rho/(tb)$.
The Lévy measure is

$$v\,(dy) = \frac{1}{\sqrt{2\pi\rho y^3}} e^{-(1/2\rho)b^2 y}\,dy. \tag{2.5}$$

We stress that both the Gamma and the inverse Gaussian process have a Lévy measure that gives much weight to small jumps and that decreases fast at infinity.
- The α-stable process, $0 < \alpha < 1$, whose Laplace transform is

$$M_t(\theta) = e^{-c\theta^\alpha t}.$$

For these processes an explicit expression for the density does not exist, unless $\alpha = \frac{1}{2}$. In this case, for $c = \sqrt{2}$, the density is, for $x > 0$,

$$f(x) = \frac{t}{(2\pi)^{1/2} x^{3/2}} e^{-t^2/2x}.$$

Notice that as $x \to +\infty$ the density decreases very slowly and, in particular, both expectation and variance are infinite. Also the Lévy measure decreases very slowly at infinity, therefore allowing very large jumps. We shall see that this model is not well suited for the growth phenomena of interest and we mention it only for reasons of completeness.

To the previous examples one should add the deterministic increasing process: $Z_t = \gamma t$. Recall also that the sum of independent subordinators is a subordinator itself.

3. Modelling by time change

A first growth model using subordinators can be obtained by time change. We simply consider that the size of the fish X_t at

time t is given by

$$X_t = Z_{A_t},$$

where Z is the subordinator and A is the von Bertalanffy type function

$$A_t = \left(1 - e^{-k(t+t_0)}\right).$$

If the subordinator Z_t has finite mean, then for some real number μ

$$E(X_t) = \mu\left(1 - e^{-k(t+t_0)}\right)$$

and the means follow therefore a VBGF, where μ plays the role of L_∞ (recall that $\mu = E(Z_1)$).

If in particular Z is a Gamma process, then $\mu = \alpha/\lambda$. It is also immediate that, if Z_t has finite variance,

$$Var(X_t) = \sigma^2\left(1 - e^{-k(t+t_0)}\right) \tag{3.1}$$

for some $\sigma^2 > 0$ ($\sigma^2 = \alpha/\lambda^2$ if Z is a Gamma process). Recall that the variance is always increasing in t. This fact makes that the models obtained by time change of a subordinator are rather rigid and are in particular unable to account for a non-monotonic behaviour of the variance, as is often observed in data.

4. Modelling by Doléans exponential

A more interesting model of growth is given by the process $(X_t)_t$ which is obtained as the solution of the stochastic equation

$$\begin{aligned} dX_t &= (L_\infty - X_{t-})dZ_t, \\ X_{-t_0} &= 0, \end{aligned} \tag{4.1}$$

where $(Z_t)_t$ is a subordinator. Here $-t_0$ denotes the time of conception, at which the size must be considered equal to 0. If the subordinator Z has a drift γ equal to 0, the solution of Eq. (4.1) is a process that remains constant between the jumps of Z, whereas it has an increment of size

$$(L_\infty - X_{t-})\Delta Z_t$$

whenever Z has a jump of size ΔZ_t. The quantity X_{t-} denotes the value of X just before the jump that occurs at time t. If we define $Y_t = L_\infty - X_t$, then Y is the solution of

$$\begin{aligned} dY_t &= -Y_{t-}dZ_t \\ Y_{-t_0} &= L_\infty. \end{aligned} \tag{4.2}$$

Therefore Y is equal to the Doléans exponential of the Lévy process $-Z$ multiplied by L_∞. The Doléans exponential of the Lévy process is a subject that has received much attention and it is possible to derive an explicit solution of Eq. (4.2), at least if one makes the assumption that the process Z cannot make jumps larger than 1 (which is a quite reasonable assumption, in our case). Under this assumption the solution of Eq. (4.2) is

$$Y_t = L_\infty e^{-\widetilde{Z}_{t+t_0}}, \tag{4.3}$$

where $\widetilde{Z}$ is another subordinator whose Lévy measure $\widetilde{\nu}$ and drift $\widetilde{\gamma}$ are obtained from ν and γ in an explicit way. More precisely $\widetilde{\nu}$ is the image of ν through the application $x \to \log(1+x)$ and $\widetilde{\gamma} = \gamma$. The solution of Eq. (4.1) is therefore

$$X_t = L_\infty(1 - e^{-\widetilde{Z}_{t+t_0}}). \tag{4.4}$$

It is easy to compute the mean $E(X_t)$, as the quantity $E(e^{-\widetilde{Z}_{t+t_0}})$ is the Laplace transform at $\theta = 1$ of the r.v. Z_{t+t_0}, which is easy to compute thanks to Eqs. (2.1) and (2.2):

$$E(X_t) = L_\infty(1 - e^{-(t+t_0)\widetilde{\phi}(1)}). \tag{4.5}$$

Hence also for this model whatever the subordinator that is chosen (provided its Lévy measure vanishes outside [0,1]), the mean of the solution of Eq. (4.1) follows a von Bertalanffy-type equation.

Note 4.1. Recall that the VBGF (Eq. (1.1)) is the solution of Eq. (4.1) when Z is the deterministic subordinator $Z_t = kt$. Therefore Eq. (4.1) appears as a natural stochastic extension of the VBGF.

Remark also that some models already present in the literature are particular instances of Eq. (4.1). It is the case of model (3.4) of Lv and Pitchford (2007) where the authors actually consider $Z_t = rt + \sigma B_t$. We point out again that the solution to Eq. (4.1) in this case does not have paths which are monotonically increasing.

The solution (Eq. (4.4)), however, has a drawback as a growth model because $X_t \to L_\infty$ as $t \to +\infty$ (unless $Z_t \equiv 0$). This implies that $Var(X_t) \to 0$ as $t \to +\infty$, which is not realistic, as it would imply that all individuals should reach the same limiting size as t increases. In order to overcome this problem it seems natural to assume that the extremal size L_∞ is itself a random variable, thus accounting for the individual variability. In this way the two sources of randomness, L_∞ and $(Z_t)_t$, appearing in Eq. (4.1) have the meaning of modelling the random individual variability and the environmental randomness, respectively. It is therefore natural to assume that L_∞ and $(Z_t)_t$ are independent. In this case the formula (Eq. (4.5)) remains valid if one replaces L_∞ by $E(L_\infty)$. Therefore the expectation $E(X_t)$ still follows a von Bertalanffy-type pattern as in Eq. (4.5), but with L_∞ replaced by $E(L_\infty)$. Similarly it is easy to compute the variance which is given by

$$\begin{aligned} Var(X_t) &= Var(1 - e^{-\widetilde{Z}_{t+t_0}})E[L_\infty^2] + Var(L_\infty^2)E[1 - e^{-\widetilde{Z}_{t+t_0}}]^2 \\ &= (e^{-(t+t_0)\widetilde{\phi}(2)} - e^{-2(t+t_0)\widetilde{\phi}(1)})E[L_\infty^2] \\ &\quad + Var(L_\infty^2)(1 - e^{-(t+t_0)\widetilde{\phi}(1)})^2. \end{aligned} \tag{4.6}$$

Notice that $Var(X_t)$ appears as the sum of two terms. As both $\widetilde{\phi}(1)$ and $\widetilde{\phi}(2)$ are positive, the first term vanishes at $-t_0$ and at $+\infty$ and it is easy to see that (unless $Z_t \equiv 0$) it increases at first and then decreases. The second one conversely is increasing in t. Therefore the variance of X_t, according to different values of the parameters, can exhibit two possible behaviours:

- increasing or
- first increasing and then decreasing.
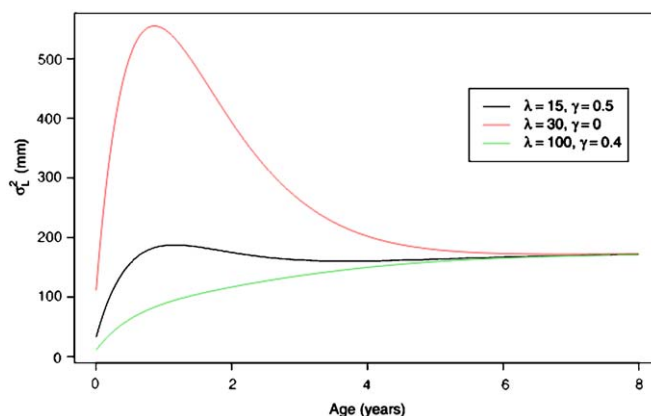
See Fig. 2 for an illustration of this typical behaviour.

Fig. 2. Behaviour of the variance (Eq. (4.10)) for $\alpha = 10$, $t_0 = 0.05$ and various values of λ and γ. Here we assumed that L_∞ is a random variable, independent of Z, with mean $= 300$ and variance $= 60$. Note that the trend can be both monotonic or exhibit one bump.

The non-monotonic behaviour of the variance is an interesting feature that is enjoyed also by some of the models proposed by Gudmundsson (2005) and Lv and Pitchford (2007). This is not surprising as the computation above only makes use of the fact that $(X_t)_t$ is the solution of Eq. (4.1) with a driving process $(Z_t)_t$ which is a Lévy process, possibly a Brownian motion as it happens to be the case for the models of the authors above.

Note 4.2. In Section 6 we adapt the models of this section and of Section 3 to a population of herrings. It is fair, however, to point out a limitation that arises when trying to model real populations with processes driven by subordinators. Recall that the driving subordinator is meant to model the randomness of the growth process arising from the environment. The assumption of independence and stationarity of the increments of the subordinator implies the assumption that the environment remains stable and stationary in time, which is a feature that can be expected to hold in real world only for a short span of time, as the effect of season and of other sources of modification of the natural habitat should introduce a time-dependent effect. It is clear, however, that more realistic models would be far more complicated and, possibly, intractable in practice. The same observation, by the way, applies to the VBGF, which is an equation that is derived under the assumption of stationarity of the environment.

In the rest of this section we give more details in two particular cases, making assumptions on Z that imply that $\widetilde{Z}$ is either a Gamma or an inverse Gaussian process.

Assume that the Lévy measure v of Z has a density h with respect to the Lebesgue measure. h must be $\geqslant 0$, must vanish on $[1, +\infty[$ and be such that

$$\int_0^1 y h(y)\, dy < +\infty.$$

Then $-Z$ has a drift $= -\gamma$ and a Lévy measure $\bar{v}$ given by the density

$$\bar{h}(y) = h(-y).$$

The image of the measure $\bar{h}(y)\, dy$ by $y \to \log(1 + y)$ is

$$\widetilde{h}(y) = e^{-y} h(1 - e^{-y}).$$

Example 4.1. (The Gamma process). Let us assume that

$$h(y) = \frac{\alpha}{-\log(1-y)}(1-y)^{\lambda-1} \tag{4.7}$$

for $0 < y \leqslant 1$ and $h(y) = 0$ for $y > 1$, where $\alpha > 0$, $\lambda > 0$. With this choice we obtain

$$\widetilde{h}(y) = \frac{\alpha}{y} e^{-\lambda y},$$

which is the Lévy measure of a Gamma process. Hence if we choose a driving subordinator Z with a Lévy measure as in Eq. (4.7) and drift γ, the solution of Eq. (4.1) is

$$X_t = L_\infty(1 - e^{-\gamma t - \widetilde{Z}_{t+t_0}}), \tag{4.8}$$

where $\widetilde{Z}$ is a Gamma process with parameters α and λ and drift γ. Thanks to Eqs. (4.1) and (2.4), for this model the mean is equal to

$$E(X_t) = E(L_\infty)\left(1 - e^{-\gamma(t+t_0)}\left(\frac{\lambda}{\lambda+1}\right)^{\alpha(t+t_0)}\right), \tag{4.9}$$

which, as already noted, is a von Bertalanffy equation with $k = \gamma + \alpha \log((\lambda+1)/\lambda)$. The variance of X_t is also easily computed

using Eq. (4.6):

$$\mathrm{Var}(X_t) = E[L_\infty^2] e^{-2\gamma(t+t_0)}\left[\left(\frac{\lambda}{\lambda+2}\right)^{\alpha(t+t_0)} - \left(\frac{\lambda}{\lambda+1}\right)^{2\alpha(t+t_0)}\right]$$
$$+ \mathrm{Var}(L_\infty)\left[1 - e^{\gamma(t+t_0)}\left(\frac{\lambda}{\lambda+1}\right)^{\alpha(t+t_0)}\right]^2. \tag{4.10}$$

It is also possible to compute the density of X_t, conditioned to the value of $L_\infty = l$, which is given by

$$f(x) = \frac{\lambda^{\alpha(t+t_0)}(l-x)^{\lambda-1}}{\Gamma(\alpha(t+t_0))l^\lambda}\left(-\log\left(1 - \frac{x}{l}\right) - \gamma(t+t_0)\right)^{\alpha(t+t_0)-1} e^{\lambda\gamma(t+t_0)} \tag{4.11}$$

for $l(1 - e^{-\gamma(t+t_0)}) \leqslant x \leqslant l$ and $f(x) = 0$ otherwise. The graph of this density for different values of t is produced in Fig. 3. See in Fig. 4 the graph of some simulated paths of $(X_t)_t$ with $\gamma = 0$.

Example 4.2. (The inverse Gaussian process). Let us assume that Z is a Levy process such that $\widetilde{Z}$ is an inverse Gaussian process with drift γ. If the Lévy measure of Z has a density

$$h(x) = \frac{1}{\sqrt{2\pi\rho}(-\log(1-x))^{3/2}}(1-x)^{(b^2/2\rho)-1}$$

for $0 < x < 1$ and $h(x) = 0$ elsewhere, then it turns out that $\widetilde{Z}$ is an inverse Gaussian process with parameters b and ρ. The density of $X_t = L_\infty(1 - e^{-Z_t - \gamma t})$ given $L_\infty = l$ is straightforward to compute and is

$$f_t(x) = \frac{t\, e^{bt/\rho}}{\sqrt{2\pi\rho}}\left(-\log\left(1 - \frac{x}{l}\right)\right)^{-3/2}\left(1 - \frac{x}{l}\right)^{(b^2/2\rho)-1} \exp\left(\frac{t^2\rho}{2\log\left(1 - \frac{x}{l}\right)}\right)$$

for $0 < x < l$ and $f_t(x) = 0$ otherwise.

One could also think of $\widetilde{Z}$ as a stable process with exponent $\alpha = \frac{1}{2}$. This choice, however, does not seem really suitable. Indeed in this case it is also possible to compute the density, as the same type of computation produces in this case the density, conditional on $L_\infty = l$,

$$g_t(x) = \frac{t+t_0}{\sqrt{2\pi}(l-x)\left(\log\frac{l}{l-x}\right)^{3/2}} \exp\left(-\frac{(t+t_0)^2}{2\log\frac{l}{l-x}}\right), \quad 0 < x < l$$
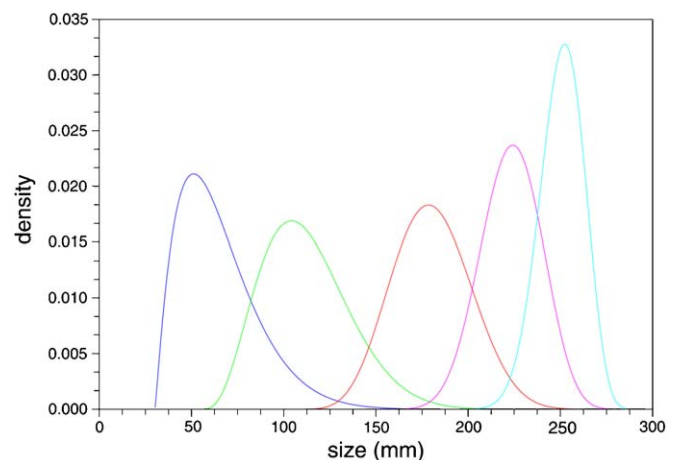

Fig. 3. The graph of the density (Eq. (4.11)) for different values of the age. The parameters here are $\alpha = 1.3$, $\lambda = 15$, $\gamma = 0.1$, $t_0 = 0.072$ and $L_\infty = 297.2$.

(with $\gamma = 0$). The expression of this density suggests that the use of stable subordinators is inappropriate in order to model fish growth, as $g(x) \to +\infty$ as $x \to l-$, *for every value* of t.

The model based on the Doléans exponential (Eq. (4.1)) appears to enjoy many interesting features. It is fair, however, to point out that in order to take advantage of it there remains the question of determining an appropriate distribution for the maximal length L_∞. It also appears difficult to find a distribution such that the density distribution of the length of the individuals of a given time t, $L_\infty(1 - e^{-Z_{t+t_0}})$ has an explicit analytic expression. This density can be computed numerically, but it may be impossible to use statistical methods based on maximum likelihood.

The question of determining a good candidate of distribution has already been tackled in Sainsbury (1980), where the author suggested a Gamma distribution.

5. The dataset

The data used in this study were collected as part of the EU FP5 project "HERGEN" (Mariani et al., 2005; Ruzzante et al., 2006), and include a homogeneous collection of genetically and phenotypically characterized North Sea Autumn Spawning herring (*Clupea harengus* L., 1758), from ICES areas zones IVa, IVb and IVc. Herring is the dominant converter of zooplankton production, using the biomass of copepods, mysids, euphausids in the pelagic environment of the northern hemisphere (Winters and Wheeler, 1987). Additionally, herring is a central prey item for higher trophic levels. The spawning period of this species in the western North Sea is September/October.
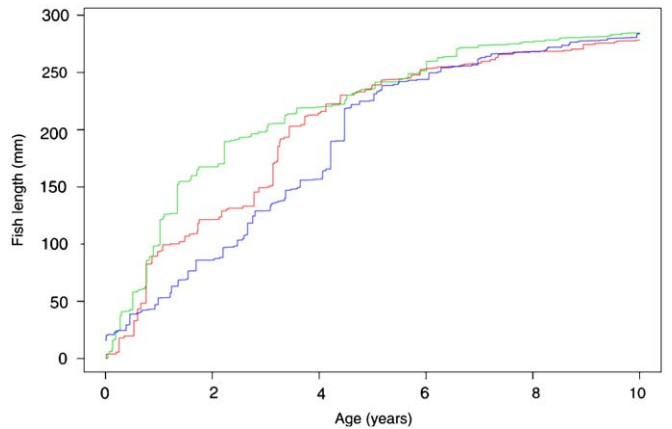

Fig. 4. A sample of paths of the solutions of Eq. (4.1), with a $\widetilde{Z}$ which is a Gamma process with $\alpha = 4$, $\lambda = 12$, $\gamma = 0$, $t_0 = 0.072$, $L_\infty = 297.2$ (deterministic) and $k = 0.59$. The process of growth is decomposed in stable step (corresponding to no growth) followed by jump of variable intensities. It is quite apparent that the variance is largest for intermediate values of time.

Table 1
Summary of the ages and abundances of the herring dataset.

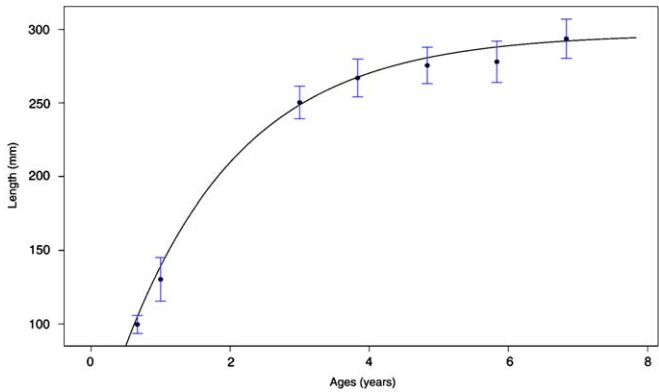| Cohort	Age (years)	No. of specimens
1	0.66	92
2	1	186
3	3	369
4	3.8	254
5	4.8	300
6	5.8	28
7	6.8	26

Fig. 5. Means and error bars for the seven cohorts with the fitted von Bertalanffy function. Notice that the curve overestimates the mean of the fifth and sixth cohorts.

The samples were collected east of the Shetland Islands in July 2005. At that moment, fish were located in that area to feed (summer feeding aggregation). Aging was performed by counting the otolith (sagitta) winter rings, following standard procedures (Ruzzante et al., 2006; AA. VV, 2007). This method was validated and tested for reliability following the procedure reported in Beverton et al. (2004). The dataset is composed by 1255 specimens belonging to seven cohorts. The abundances and the ages of these cohorts are shown in Table 1. The values of the empirical means and variances of the cohorts can be found in Fig. 5.

6. Analysis of the herring dataset

In order to apply the models developed in Sections 3 and 4 to the analysis of a dataset as the herring one described in the previous section, one is confronted with two kinds of problems.

The first one is the determination of the appropriate driving subordinator. Actually subordinators form a large family, every Lévy measure on $\mathbb{R}^+$ satisfying Eq. (2.3) being associated to a corresponding subordinator. This is therefore a non-parametric problem and it appears very difficult to determine this Lévy measure starting from data in the form of cohorts, as is the case for the herring dataset.

The second order of problems comes from the fact that these models are to be considered only approximatively correct, as they do not take into account the time variability of the environment (see Note 4.2).

In this section we produce the results obtained fitting the models of Sections 3 and 4 to the herring dataset. In both cases we decided to work with the Gamma model, as it seemed to give better results than the inverse Gaussian.

In order to have a benchmark for the analysis of the dataset, we shall first adapt a normal Gaussian model inside each cohort. This is the default choice in the literature (Imsland et al., 1998). Given the age t_i, the length X_{t_i} is assumed to be normally distributed with mean $L_\infty(1 - e^{-k(t_i+t_0)})$ and variance σ_i^2 (different cohorts are allowed to have different variances). This model requires therefore a number of parameters that is equal to $3+$ the number of cohorts.

It is worth noting that this is not a mixture model, as we are able to assign every individual to its class. Recall also that, as stated in the introduction, this is not an IBM.

The normal and time change models have been estimated by means of maximum likelihood, which appears to be the most natural method for them. As already pointed out in Section 4, for the Doléans exponential model, which is the most promising, the maximum likelihood method is inapplicable, as the distribution

Table 2
Summary of the estimates obtained with the different models.

Model	Empirical	Normal benchmark	Time change	Doléans
$\hat{L}_\infty$	–	289.19	291.52	297.28
$\hat{k}$	–	0.67	0.64	0.63
$\hat{t}_0$	–	0.046	0.047	0.072
$\hat{\mu}_1$	99.59	98.22	107.39	100.88
$\hat{\mu}_2$	130.29	136.37	142.96	136.32
$\hat{\mu}_3$	250.31	249.06	250.53	244.20
$\hat{\mu}_4$	267.01	266.20	267.56	266.87
$\hat{\mu}_5$	275.55	277.41	278.94	278.63
$\hat{\mu}_6$	278.03	283.15	284.91	284.74
$\hat{\mu}_7$	293.69	286.09	288.05	287.91
$\hat{\sigma}_1^2$	37.28	37.27	80.65	172.16
$\hat{\sigma}_2^2$	220.52	220.56	107.35	186.22
$\hat{\sigma}_3^2$	123.19	123.42	188.14	162.28
$\hat{\sigma}_4^2$	164.93	164.92	200.92	160.41
$\hat{\sigma}_5^2$	152.96	153.37	209.46	162.96
$\hat{\sigma}_6^2$	197.96	198.05	213.95	166.52
$\hat{\sigma}_7^2$	177.21	177.21	216.31	169.41

of the maximal size L_∞ is not known. We therefore resorted to the least squares method in order to fit the moments.

The estimates obtained with different models are shown in Table 2.

It is apparent that the normal model has a good agreement with the data, which is not surprising, given also the availability of many parameters.

6.1. Time change model

As already hinted in Section 3 this kind of models shows a certain rigidity. We give the estimates obtained using it for completeness sake only. Assuming that the lengths at time t_i follow a $\Gamma(\alpha(1 - e^{-k(t_i+t0)}), \lambda)$ distribution, we have a model with four parameters (whatever the number of cohorts). The estimated values are $\alpha = 388.2$, $\lambda = 1.33$, $k = 0.64$ and $t_0 = 0.047$. From this we deduce $\hat{L}_\infty$ as the ratio $\hat{\alpha}/\hat{\lambda}$ (see Table 2). Estimates of the means and variances are obtained using Eq. (3.1).

Notice that the estimated means are very close to the empirical means (with the exception of cohort 1). However, it is clear that this model cannot accommodate the variances. Recall that for this model the variances are necessarily increasing with time, at a difference with the behaviour of the empirical variances.

As already mentioned above, the model based on time change is a particularly parsimonious explanation of the data in terms of the number of parameters, but has limited capacities of catching some relevant features.

6.2. Doléans exponential

We tried a model based on the solution of Eq. (4.1) with respect to a Gamma process, as in Example 4.1.

According to this model, the density of the observations follows a distribution, given $L_\infty = l$, that is given in Eq. (4.11). Noting that the mean values of the lengths of the different cohorts must lie on the VBGF, we decided to perform a two-stage method of moments, in which, at first, the empirical means are interpolated by a VBGF function (Eq. (1.1)), therefore estimating the parameters $E(L_\infty)$, k and t_0. In the second stage, we interpolated the empirical variances with the function (Eq. (4.10)). In this way we obtained estimates of α, λ, γ together with an estimate of $\sigma_{L_\infty}^2 = \text{Var}(L_\infty)$. Recall that the quantities k and α, λ, γ are related by the constraint $k = \gamma + \alpha \log((\lambda + 1)/\lambda)$. For both

stages we used a least squares method in order to fit the moments.

See Fig. 5 for the graph of the fitted von Bertalanffy function and the empirical means with error bars.

This method produces an estimate of the mean and variance of L_∞ without assuming any distribution for L_∞. The estimates are $\alpha = 1.3$, $\lambda = 15$, $\gamma = 0.55$, $k = 0.63$, $t_0 = 0.072$, $E(L_\infty) = 297.3$ and $\sigma_{L_\infty}^2 = 160$. Finally, in order to obtain a concrete and visually appreciable estimate of the density of each cohort, we computed it numerically assuming for L_∞ a Gamma distribution having parameters matching the estimated mean and variance, which is with parameters $\alpha_0 = 552.34$ and $\lambda_0 = 1.86$. Fig. 6 produces a
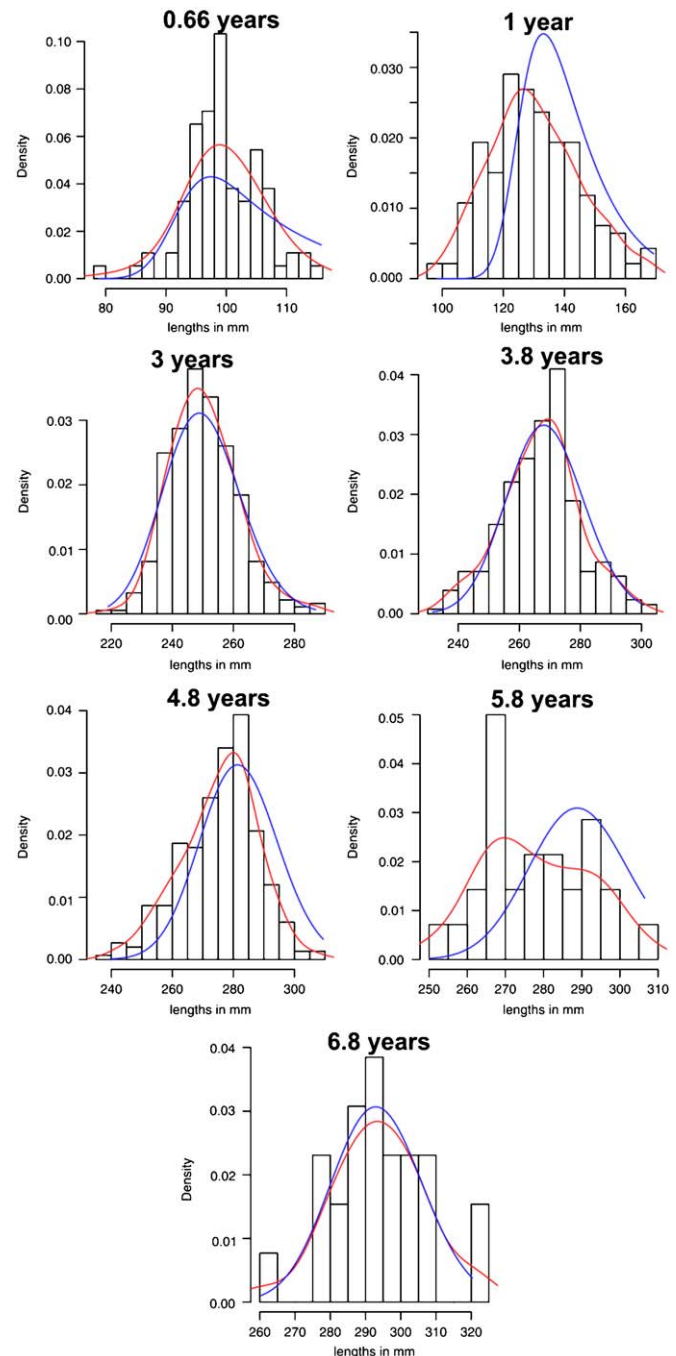

Fig. 6. Kernel estimate (in red) and Doléans exponential model (based on a Gamma subordinator) (in blue) densities for the different cohorts. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

comparison between the density obtained as described above from the model and an empirical density, produced from the data with a usual kernel estimator. Taking into account the limited number of parameters employed by this model one can appreciate the nice fit for some of the cohorts. Discrepancies can be found for the first two cohorts and for the sixth one. For the fifth one there is an evident discrepancy: notice, however, that, as pointed out in Fig. 5, also the estimated VBFG curve that fits the mean of this fifth cohort. At this point, one should keep Note 4.2 in mind.

7. Discussion

Although several stochastic models are available in the literature in order to compute the time evolution of the distribution of the size of a population, a certain number of drawbacks still remain, leading to the need of a more appropriate formulation that is able to take into account some key aspects of animal growth.

These aspects basically concern the inclusion of the different sources of variability in growth rate among individuals, which directly determine the size density observed at different ages, and the property of the growth process of being increased. In this work we developed two models based on the use of subordinators as driving processes. Our aim was to provide a new stochastic model, mainly of theoretical interest for the description of the growth process for a large class of organisms. It turns out that one of them, the model based on the Doléans exponential giving rise to a Gamma process, is also able to produce an appreciable fit with the data. In particular it is able to reproduce the non-monotonic trend of the variance, which is the measure of intra-cohort variability of size. This is an important feature: a reduction of the demographic variance starting at some time is often observed in data (e.g. Gudmundsson, 2005).

In general, the variability in individual size increases within the same cohort through time when the individual growth rate is positively correlated with itself in time (Gudmundsson, 2005). This phenomenon is called "growth depensation".

Conversely, a reduction of the cohort variance (the so-called "growth compensation") has been up to now explained with the fact that the survival of individuals is not independent and identically distributed (Kendall and Fox, 2002). In other words, there needs to be some systematic structure in the population. Common biological mechanisms that can produce a reduction of demographic variance are contest competition (including territoriality), long-lived individuals with lifetime demographic traits ("individual heterogeneity"), maternally imposed variation, and directional or balancing selection.

Observe, however, that our model explains this typical behaviour of the variance without introducing an explicit selection mechanism related to some size-dependent mortality. By this we do not claim that such a mechanism does not exist (and we think that it might be natural to devise a more sophisticated model including this feature), but simply that it is not necessary in order to explain the observed behaviour of the empirical variances.

The possibility to model and predict this aspect of fish growth seems to be of great importance in both theoretical studies and fishery management applications. This is particularly true if referred to the increasing use of measurement of the growth pattern of organisms (like fish) as an ecological indicator (Bennett et al., 1995). Further, the assessment of growth pattern, combined with other measurement of physiological condition, has the potential to yield information on the history of environmental stress (e.g. from contaminants) or selective pressures (e.g. fisheries) and adaptation to environmental change (Jorgensen et al., 2007). In fact, the statistical analyses of long-term data from

some exploited fish stocks have revealed evolutionary changes in reaction norms (Ernande et al., 2004). Accuracy of the growth model at the basis of these analyses obviously affects the results, so that the development of sound approach to model and predict size distribution is the main target.

A future direction will concern the assessment of the effect of the aging error on the estimates obtained using our model. Even if the aging method is reliable, it seems that errors in aging young cohorts can significantly affect the estimates. This is due to the fact that both VBGF and the variance function converge with age.

The present work represents a first attempt to apply the Lévy processes to the subject. We think that these and the ideas developed in this paper can be adapted to introduce stochasticity into more complex models of growth (see Edwards et al., 2007).

8. Conclusion

The aim of the present work was the development of a stochastic model of growth, and therefore an IBM for the distribution of data, that is compatible with the von Bertalanffy function and might be able to catch more of the features empirically observed.

The model developed in relation with the Doléans exponential shows interesting features. As other models already in the literature (Wang, 1999; Gudmundsson, 2005; Lv and Pitchford, 2007) it is able to explain the observed non-monotonic behaviour of the variances. However, it is the first stochastic model producing paths that are increasing, thus giving a realistic random model of individual growth.

Our main objective was mainly theoretical, but we think that the models developed in Section 4 can be of interest for practical applications such as stock assessment and forecasting. In this direction, however, some questions require a deeper investigation. These are

- the determination (elicitation) of a suitable distribution for the maximal size L_∞;
- the determination of the appropriate subordinator.

References

AA., VV., 2007. Report of Herring Assessment Working Group for the Area South of 628 N. ICES CM 2007/ACFM: 11.

Arino, O., Shin, Y., Mullon, C., 2004. A mathematical derivation of size spectra in fish populations. C.R. Biol. 327, 245–254.

Bennett, W., Ostrach, D., Hinton, D., 1995. Condition of larval striped bass in a draught-stricken estuary: evaluating pelagic food web limitation. Ecol. Appl. 5, 680–692.

Beverton, R., Hylen, A., Østvedt, O.-J., Alvsvaag, J., Iles, T., 2004. Growth, maturation, and longevity of maturation cohorts of Norwegian spring-spawning herring. ICES J. Mar. Sci. 61, 165–175.

Brett, J., 1979. Environmental factors and growth. In: Fish Physiology. Bioenergetics and Growth, vol. III. Academic Press Inc., New York.

Caswell, H., 2001. Matrix Population Models: Construction, Analysis, and Interpretation. Sinauder Associated, Sunderland, Massachusetts, USA.

Conover, D., Munch, S., 2002. Sustaining fisheries yields over evolutionary time scales. Science 297 (5578), 94–96.

Cont, R., Tankov, P., 2004. Financial Modelling with Jump Processes. Chapman & Hall/CRC Financial Mathematics Series, Chapman & Hall/CRC, Boca Raton, FL.

DeAngelis, D., Rose, K., Crowder, L., Marshall, E., Lika, D., 1993. Fish cohort dynamics: application of complementary modelling approaches. Am. Nat. 142 (4), 604–622.

Edwards, A., Phillips, R., Watkins, N., Freeman, M., Murphy, E., Afanasyev, V., Buldyrev, S., da Luz, M., Raposo, E., Stanley, H., 2007. Revisiting Lévy flight search patterns of wandering albatrosses, bumblebees and deer. Nature 449, 1044–1049.

Ernande, B., Dieckmann, U., Heino, M., 2004. Adaptive changes in harvested populations: plasticity and evolution of age and size at maturation. Proc. R. Soc. Lond. 271 (1537), 415–423.

Essington, T., Kitchell, J., Walters, C., 2001. The von Bertalanffy growth function, bioenergetics, and the consumption rates of fish. Can. J. Fish. Aquat. Sci. 58, 2129–2138.

Fujiwara, M., Kendall, B., Nisbet, R., 2004. Growth autocorrelation and animal size variation. Ecol. Lett. 7, 106–113.

Gudmundsson, G., 2005. Stochastic growth. Can. J. Fish. Aquat. Sci. 62, 1746–1755.

Hinrichsen, H.-H., Buehler, V., Clemmesen, C., 2007. An individual-based model for direct conversion of otolith into somatic growth rates. Fish. Oceanogr. 16 (3), 207–215.

Imsland, A., Jonassen, T., Langston, A., Hoare, R., Wergeland, H., FitzGerald, R., Mulcahy, M., Stefansson, S., 2002. The interrelation of growth performance and disease resistance of juvenile Atlantic halibut (*hippoglossus hippoglossus* L.) from different populations. Aquaculture 204, 167–177.

Imsland, A., Nilsen, T., Folkvord, A., 1998. Stochastic simulation of size variation in turbot: possible causes analysed with an individual based-model. J. Fish. Biol. 53, 237–258.

Jorgensen, C., Enberg, K., Dunlop, E., Arlinghaus, R., Boukal, S., Brander, K., Ernande, K., Gardmark, A., Johnston, F., Matsumura, S., Pardoe, H., Raab, K., Silva, A., Vainikka, A., Dieckmann, U., Heino, M., Rijnsdorp, A., 2007. Managing evolving fish stocks. Science 318, 1247–1248.

Karlin, S., Taylor, H.M., 1981. A Second Course in Stochastic Processes. Academic Press Inc. (Harcourt Brace Jovanovich Publishers), New York.

Kendall, B., Fox, G., 2002. Variation among individuals and reduced demographic stochasticity. Cons. Biol. 16 (1), 109–116.

Kooijman, S., 2000. Dynamic Energy and Mass Budgets in Biological Systems, second ed. Cambridge University Press, Cambridge, UK.

Lv, Q., Pitchford, J.W., 2007. Stochastic von Bertalanffy models, with applications to fish recruitment. J. Theoret. Biol. 244 (4), 640–655.

Mariani, S., Hutchinson, W., Hatfield, E.M.C., Ruzzante, D., Simmonds, E.J., Dahlgren, T.A.C., Brigham, J., Torstensen, E., Carvalho, G., 2005. Biocomplexity in a highly migratory marine pelagic fish, Atlantic herring. Mar. Ecol. Progr. Ser. 303, 245–257.

Mulligan, T., Leaman, B., 1992. Length-at-age analysis: can you get what you see? Can. J. Fish. Aquat. Sci. 49, 632–643.

Rilling, G., Houde, E., 1999. Regional and temporal variability in the distribution and abundance of bay anchovy (*anchoa mitchilli*) eggs, larvae, and adult biomass in the Chesapeake bay. Estuaries 22, 1096–1109.

Ruzzante, D., Mariani, S., Bekkevold, D., Andre, C., Mosegaard, H., Clausen, L., Dahlgren, T., Hutchinson, W., Hatfield, E., Torstensen, E., Brigham, J., Simmonds, E., Laikre, L., Larsson, L., Stet, R., Ryman, N., Carvalho, G., 2006. Biocomplexity in a highly migratory marine pelagic fish, Atlantic herring. Proc. R. Soc. Lond. 273, 1459–1464.

Sainsbury, K., 1980. Effect of individual variability on the von Bertalanffy growth equation. Can. J. Fish. Aquat. Sci. 37, 241–247.

Sato, K.-I., 1999. Lévy Processes and Infinitely Divisible Distributions. Cambridge Studies in Advanced Mathematics, vol. 68. Cambridge University Press, Cambridge (Translated from the 1990 Japanese original, revised by the author).

Sumpter, J., 1992. Control of growth of rainbow trout (*oncorhynchus mykiss*). Aquaculture 100, 299–320.

Uchmanski, J., 2000. Individual variability and population regulation: an individual-based model. Oikos 9, 539–548.

von Bertalanffy, L., 1938. A quantitative theory of organic growth. Human Biol. 10, 181–213.

von Bertalanffy, L., 1957. Quantitative laws in metabolism and growth. Quart. Rev. Biol. 32 (3), 217–231.

Wang, Y., 1999. Estimating equations for parameters in stochastic growth models from tag-recapture data. Biometrics 55, 900–903.

Wang, Y., Thomas, M., 2003. Accounting for individual variability in the von Bertalanffy growth model. Can. J. Fish. Aquat. Sci. 52, 1368–1375.

Weatherley, A., Gill, H., 1987. The Biology of Fish Growth. Academic Press Inc. (Harcourt Brace Jovanovich Publishers), New York.

Winters, G., Wheeler, J., 1987. Clupeoid fishes of the world (suborder clupeioidei). An annotated and illustrated catalogue of the herrings, sardines, pilchards, sprats, shads, anchovies and wolf-herrings. Can. J. Fish. Aquat. Sci. 44, 882–900.