# An individual-based approach to year-class strength estimation 

S Thanassekos<br>Virginia Institute of Marine Science<br>RJ Latour<br>Virginia Institute of Marine Science<br>Mary C. Fabrizio<br>Virginia Institute of Marine Science

Follow this and additional works at: https://scholarworks.wm.edu/vimsarticles
Part of the Aquaculture and Fisheries Commons

## Recommended Citation

Thanassekos, S; Latour, RJ; and Fabrizio, Mary C., "An individual-based approach to year-class strength estimation" (2016). VIMS Articles. 794.
https://scholarworks.wm.edu/vimsarticles/794

This Article is brought to you for free and open access by the Virginia Institute of Marine Science at W\&M ScholarWorks. It has been accepted for inclusion in VIMS Articles by an authorized administrator of W\&M ScholarWorks. For more information, please contact scholarworks@wm.edu.

# Original Article <br> An individual-based approach to year-class strength estimation 

Stéphane Thanassekos*, Robert J. Latour, and Mary C. Fabrizio<br>Virginia Institute of Marine Science, College of William \& Mary, PO Box 1346, Gloucester Point, VA 23062, USA<br>*Corresponding author: tel: +1 804684 7363; fax: +1 804684 7327; e-mail: thanassekos@vims.edu.<br>Thanassekos, S., Latour, R.J., and Fabrizio, M. C. An individual-based approach to year-class strength estimation. - ICES Journal of Marine Science, 73: 2252-2266.

Received 22 December 2015; revised 22 March 2016; accepted 25 March 2016; advance access publication 21 April 2016.


#### Abstract

Estimating year-class strength-the number of larvae hatched in a given year-from survey data is key to investigating fish population dynamics. Year-class strength can be estimated from catch-at-age data using catch curves. In practice, most catch-curve assumptions are violated, which can result in spurious estimates of year-class strength. Among the simplifying assumptions is that pooling individuals into annual age-classes provides a representation of the population age structure that is adequate for estimating mortality. This oversimplification is unnecessary when age data are available at finer scales, and can lead to biased results. We present a new method to estimate past year-class strength with a set of equations that apply to each sampled individual. Through the reconstruction of individual histories from hatch to capture, this approach takes full advantage of the individual resolution of survey data, enables the incorporation of the processes that violate catch-curve assumptions, and provides more accurate year-class strength estimates.


Keywords: catch curve, individual-based, mortality, selectivity, survey data, year-class strength.

## Introduction

Year-class strength is broadly defined as the number of fish spawned or hatched in a given year (Ricker, 1975). Estimating year-class strength from field samples is key to investigating the environmental drivers of fish abundance (Koonce et al., 1977; Stevens, 1977), evaluating the effectiveness of management strategies (Catalano et al., 2009), interpreting trends in past and future catches (Gudmundsdottir et al., 2007; Planque et al., 2012; Nagid et al., 2015), and assessing potential effects of global change on fish populations (Ottersen and Loeng, 2000).

The simplest method for inferring year-class strength relies on the application of the catch-curve equation (Baranov, 1918). According to this equation, the number of individuals in a population decreases exponentially as a function of age, provided a series of assumptions are met, including the assumption of constant mortality. Age-specific catches must be proportionally related to abundance and this relationship must be constant over time. Thus, information on the number of captures at age allows the initial number of age- 0 individuals in the population to be backcalculated. This mathematical reconstruction is the starting point of an array of methods that have enabled estimation of year-class strength from juvenile fish surveys (Cowx and Frear, 2004) and
reconstruction of daily hatch-date frequency distributions from larval and juvenile fish surveys (Methot, 1983; Campana and Jones, 1992; Fortier and Quiñonez-Velazquez, 1998).

Catch-curve analyses further enable the estimation of the total mortality rate $(Z)$ in a population; in such applications, $Z$ is estimated by the exponential rate of decrease in age frequency data (Robson and Chapman, 1961). As such, catch curves are instrumental tools for assessing stock status (e.g. Thorson and Prager, 2011) and informing harvest strategies (Wayte and Klaer, 2010) from a limited amount of data. For a population that is sampled in a single year, each age corresponds to a different year-class (with earlier year-classes being sampled at older ages), and mortality estimates may be valid only under the assumption of constant year-class strength (Robson and Chapman, 1961). Alternatively, departures from the exponential decrease in catch-at-age may be interpreted as a signal of changing year-class strength, which forms the basis of the catch-curve residuals approach (Maceina, 1997; Catalano et al., 2009). For fisheries-independent surveys that are repeated over multiple years, year-classes may be observed at several ages and analysed separately (Pope, 1972; Ricker, 1975; Serns, 1986). This approach to year-class strength estimation is referred to as the cohort method (Tetzlaff et al., 2011) or the longitudinal catchcurve method (Tuckey et al., 2007).

The assumptions of the catch-curve approach are: (i) mortality is constant among ages and over time, (ii) selectivity of the sampling gear is constant over time, (iii) fish availability is constant in space and time, (iv) catch is proportional to abundance, (v) in the case of subsampling, subsamples are proportional to the survey catch and constant over time, (vi) there is no emigration/immigration occurring in the population, and (vii) individuals are aged without error. In essence, after pooling sampled individuals by year-class, an estimate of year-class strength is given by the $y$-intercept of a linear regression of log-transformed counts-at-age.

In reality, most of the assumptions underlying the catch-curve approach are violated (Dunn et al., 2002; Catalano et al., 2009), thus requiring modifications of the simple approach outlined above (Xiao, 2005; Schnute and Haigh, 2007; Wayte and Klaer, 2010; Thorson and Prager, 2011). One core issue that is common to all approaches utilizing catch-at-age data is that sampled individuals are pooled into discrete (integer) age-classes. This disregards the uniqueness of individual histories from hatch to capture, generates inaccuracies, and results in computations that do not take advantage of the resolution of the available data afforded from sampled fish that are aged individually. This may also introduce bias if fish are not sampled exactly on a round number of years after their hatch date, otherwise individuals are either younger or older than their assigned age.

One numerical approach to account for the effect of individual variability on population-level processes is to apply individualbased models (IBMs) (e.g. DeAngelis and Mooij, 2005). Since their early use in fisheries science (Beyer and Laurence, 1980; DeAngelis et al., 1980), IBMs have been instrumental in clarifying ecological questions spanning a wide range of scales, from larval trophodynamics (e.g. Fiksen and MacKenzie, 2002) to ecosystem-scale species interactions (e.g. Rose et al., 2015). Drawing from the ability of IBMs to explicitly account for individual variability, an individualbased approach (IBA) is used here to extract information from simulated survey data (mimicking real-world survey data) by accounting for processes that result in the inclusion of each fish in the dataset. Because some of the capture processes vary among individuals (for example, the size-dependent gear selection of an individual), such an approach should result in more accurate computations than approaches that pool individuals within age classes (such as the catch-curve approach).

The IBA was inspired by the approach adopted by larval fish scientists for "hatch-date analysis" (Methot, 1983; Yoklavich and Bailey, 1990; Campana and Jones, 1992; Fortier and QuiñonezVelazquez, 1998). The goal of this analysis is to adjust hatch-date frequency distributions computed from field-caught larvae to account for the underrepresentation of old individuals due to mortality and dispersal. In essence, the hatch-date analysis has the same goal as year-class strength estimation, which is to derive an estimate of the initial number of larvae that later resulted in the capture of individuals, but with a daily resolution. In all methods, a certain pooling of individuals is required; by age-classes within year-classes for year-class strength estimations, or, by daily ages within hatch dates for hatch-date analyses. In this study, the initial number of larvae that resulted in the capture of each fish is estimated individually, before pooling year-classes and summing individual-based estimates. By reconstructing the series of events that resulted in the inclusion of each individual in an observational dataset, the method improves the accuracy of year-class strength estimates while avoiding the loss of information caused by prematurely pooling data.

The IBA we describe estimates year-class strength as the number of age- 0 individuals released into the environment that later resulted in the capture of each fish. This contrasts with the application of the catch-curve method which estimates the number of age-0 individuals released into the environment that later resulted in the capture of a given number of individuals across all ages. Based on the reconstruction of individual histories, from hatch to capture, our IBA takes full advantage of the individual resolution of commonly available survey data, incorporates the processes that are assumed constant for catch-curve analysis, and provides more accurate estimates of year-class strength.

To demonstrate the method, an individual-based population dynamics model was used to generate several simulated survey datasets. For illustration, the simulations were based on a hypothetical northern hemisphere, spring-hatching, moderately long-lived fish. From the simulated survey data (where the actual year-class strength is known), we estimated year-class strength using both the catchcurve method and the novel IBA and compared outputs to assess the performance of each method. We also investigated the effects of violating several of the catch-curve assumptions by changing population-level mortality rates and survey sampling regimes.

## Methods

## Individual-based population dynamics model

The population model (Figure 1) was designed to generate data mimicking real-world survey data with a monthly resolution. Over a 10-year simulation, a set number of larvae were added to the population each year around April (50\% in April, 25\% in March and May), and for each individual, ageing, growth, survival, and probability of observation by the survey were computed for each month.

## Ageing

Each individual starts its life with a biological age in months ( Age $_{\mathrm{M}}$ ) and an age-class in years $\left(\mathrm{Age}_{\mathrm{C}}\right)$ set to zero. At the end of each month, the biological age is incremented by 1 ; every January, the age-class is incremented by 1 .

## Growth

Individual growth was simulated as a deviation from a mean von Bertalanffy growth model, the parameters of which were assumed to be $\bar{k}=0.2$ and $\overline{L_{\infty}}=100 \mathrm{~cm}$. Individuals start their growth trajectory from an initial length randomly generated from a normal distribution with a mean of 0.5 cm and a standard deviation of 0.2 cm . An individual grows from its initial length to an asymptotic length $\left(L_{\infty}\right)$ randomly generated from a normal distribution with mean of $\overline{L_{\infty}}$ and a standard deviation of $10 \mathrm{~cm} . L_{\infty}$ values are generated once per individual and stored during simulations. To account for the often observed negative correlation between $k$ and $L_{\infty}, k$ was computed as a function of $L_{\infty}$ :

$$
\begin{equation*}
k=-0.00125 L_{\infty}+0.325 \tag{1}
\end{equation*}
$$

As a result, $k$ ranged from 0.15 to 0.25 over the range of $L_{\infty}$ values (ca. 140 to 60 cm ) and equalled $0.2(k=\bar{k})$ when $L_{\infty}=\overline{L_{\infty}}$. To simulate a rapid increase in length-at-age variability during the first year of life, a growth rate deviation was randomly generated from a normal distribution with mean of zero and standard deviation of 0.03 at the start of the simulation for each individual. This deviation was added to $k$ at each time-step during the first year of life (i.e. 12 months). To update the length of each individual at each time-step, a linearization of the von Bertalanffy equation was


Figure 1. Schematic of the computations occurring in each month (here, March of the 9th year) for each individual in the population dynamics model, where a set number of larvae is added to the population each year around April. In each month, growth, survival, and probability of observation by the survey are computed sequentially ( Age $_{C}$, age-class in years; Age $_{M}$, age in months). Outputs from the growth submodel are shown in inset with curves indicating the range, quartiles, and median of length-at-age. Observation implies that individual data (age-class, length, and date of capture) are recorded. Figure modified from Thanassekos et al. (2014).
used:

$$
\begin{equation*}
L_{(t)}=L_{(t-1)}+\frac{k\left(L_{\infty}-L_{(t-1)}\right)}{12} \tag{2}
\end{equation*}
$$

where $L_{(t)}$ is the length of an individual at time $t$, and the denominator of 12 allows the scaling of the annual growth to the monthly time-step of the model (see inset in Figure 1).

## Survival

Based on a given annual mortality rate $Z$ (year ${ }^{-1}$ ), the monthly probability of dying $\left(P_{\mathrm{d}}\right)$ is computed for each individual at each time-step as (e.g. Rose and Cowan, 1993):

At each time-step, an individual survives and enters the next timestep if the probability of dying is less than a random number generated from a uniform distribution bounded between 0 and 1 . Three types of instantaneous mortality rates were separately used in the model, (i) a constant natural mortality $M$ (year ${ }^{-1}$ ), (ii) an agedependent natural mortality $M_{\text {(age) }}\left(\right.$ year $\left.^{-1}\right)$, and (iii) a combined mortality $Z_{(\text {age } e t)}=M_{\text {(age) }}+F_{(t)}$, where $F_{(t)}\left(\right.$ year $\left.^{-1}\right)$ is a timevarying fishing mortality (Figure 2a and b ).

For the age-dependent mortality, a mortality rate for a given individual was assigned by computing its biological age in years ( $\mathrm{Age}_{\mathrm{Y}}$ )
as a function of its age in months $\left(\mathrm{Age}_{\mathrm{M}}\right)$ as:

$$
\begin{equation*}
\operatorname{Age}_{\mathrm{Y}}=\text { floor }\left(\frac{\text { Age }_{\mathrm{M}}}{12}\right) \tag{4}
\end{equation*}
$$

where floor $(x)$ returns the closest integer below $x$.

## Observation

In the model, observation denotes capture of an individual by the survey and subsequent processing for length and age information.


Figure 2. Parameters used in the population dynamics model. Natural mortality was either constant (set at $M=0.83$ year ${ }^{-1}$, the mean of age $0-5$ years from the age-dependent mortality) or age-dependent (a). Time-varying fishing mortality was set to arbitrary values (b). The proportion of fish taken from a given sample and aged $\left(P_{\mathrm{a}}\right)$ was set to arbitrary values and is shown here for a sampling regime where multiple samples were collected in some years from the simulated population (c). The two types of selection probabilities ( $P_{\text {sel }}$ ) used to subsample individuals were computed according to Equations (5 and 6) (d).

Occasionally during field sampling, if the catch is too large and/or ageing of fish is too costly, only a fraction of the individuals captured are aged. To include this process in the model, the proportion of fish aged $\left(P_{\mathrm{a}}\right)$ from a given sample was set as variable in some simulations (Figure 2c). If on a given sampling date $P_{\mathrm{a}}<1$, then a $P_{a}$ fraction of all individuals available for capture on that date is randomly observed. In addition, two length-dependent functions ( $P_{\text {sel }}$ ) were applied to simulate length-dependent gear selectivity as either a logistic function (i.e. asymptotic selectivity):

$$
\begin{equation*}
P_{\text {sell }}=\frac{1}{1+\mathrm{e}^{-a(\text { Length }-b)}}, \tag{5}
\end{equation*}
$$

or a double-logistic function (i.e. dome-shaped selectivity):

$$
\begin{equation*}
P_{\text {sel2 }}=\left(1-\frac{1}{1+\mathrm{e}^{-c(\text { Length }-d)}}\right) \times \frac{1}{1+\mathrm{e}^{-a(\text { Length }-b)}}, \tag{6}
\end{equation*}
$$

with parameters $a=0.5 ; b=20 ; c=0.1$; and $d=50$ (Figure 2d). An individual is selected on a given sampling date if $P_{\text {sel }}$ is greater than a random number generated from a uniform distribution bounded between 0 and 1 . During any given sampling event, not every age-class available for capture is always sampled (Figure 3). This has implications for the total number of individuals observed at a given age (Thanassekos et al., 2012) and therefore for the fit of catch-curves. To account for this process in the model, randomly chosen age-classes were excluded from survey captures (referred to as "missing age-classes") in some simulations.

Ten sampling regimes were used to subsample individuals in the model; this was accomplished by combining different sampling frequencies, proportions of individuals aged, selectivities, and the occurrence of missing age-classes in simulations (Table 1). Under the most restrictive sampling regime (SR10), an individual was included in model outputs if it (i) belonged to one of the randomly selected age-classes, (ii) was selected by the gear, and (iii) was aged on a given sampling date.

## Simulations

A set of 25 simulated datasets was generated using the population dynamics model (Table 2). These combine different mortality rates and sampling regimes along with constant or variable numbers of larvae released per year. In a given dataset, each row corresponded to an observed individual characterized by length (cm), age-class (years), year of capture (Capt, $1-10$ ), and month of capture $\left(\mathrm{Capt}_{\mathrm{M}}, 1-12\right)$. Independently of the method used to estimate year-class strength, the first step was to compute the year-class of each individual as ( $H_{\mathrm{Y}}$, the year of hatch):

$$
\begin{equation*}
H_{Y}=\text { Capt }_{Y}-\text { Age }_{C} . \tag{7}
\end{equation*}
$$

Subsequently, the catch-curve and the IBA were used to estimate year-class strength using each simulated dataset.

In addition to the 25 simulated datasets mentioned above, a set of simulations was designed to investigate the impact of variability in natural mortality on year-class strength estimates. Based on simulation S25 (Table 2), increasing levels of noise were added to the age-dependent natural mortality assigned to each individual at each time-step during simulations by generating random mortality rates from a uniform distribution bound between $\left(M_{(a g e)}-M_{(\text {age })} \times M_{\text {var }}\right)$ and $\left(M_{(\text {age })}+M_{(\text {age })} \times M_{\text {var }}\right)$, where $M_{\text {var }}$ varies between 10 and $100 \%$ and represents variability in the mortality rate.

## Year-class strength estimated using catch-curves

Using the catch-curve method, year-class strength is estimated as the $y$-intercept of the linear regression of the log-transformed number of individuals in each annual age-class. This method relies on several assumptions (see the Introduction section) but has the advantage of providing an estimate of the mortality rate as determined by the slope parameter of the regression. For a survey that is repeated annually, year-classes (individuals hatched in the same year $H_{\mathrm{Y}}$ ) are


Figure 3. Distribution of age-classes (circled numbers) from year-classes sampled on different dates (black triangles) over 10 years. When all age-classes are sampled on all sampling dates (a), the number of times a year-class has been sampled ( Ns ) corresponds to the number of samples collected after the hatch of that year-class. In reality, all age-classes are seldom sampled on each sampling date (b). The frequency of sampling shown here corresponds to sampling regimes where multiple samples are taken in some years from the simulated population (SR6-SR10, Table 1).

Table 1. Sampling regimes used to subsample individuals in the population dynamics model.

| Sampling regime | Frequency | $\mathrm{P}_{a}$ | $P_{\text {sel }}$ | Missing Age $_{\text {C }}$ |
| :---: | :---: | :---: | :---: | :---: |
| SR1 | Once in 1 of 12 months, in year 10 | 1 | 1 | No |
| SR2 | Every June | 1 | 1 | No |
| SR3 | Every June | 1 | $P_{\text {sel1 }}^{\text {a }}$ | No |
| SR4 | Every June | 1 | $P_{\text {sel2 }}^{\text {a }}$ | No |
| SR5 | June 1; February 4; September 5; March 8; December 10 | 1 | 1 | No |
| SR6 | Multiple samples in years 1; 4; 5; 8; 9; 10 | 1 | 1 | No ${ }^{\text {b }}$ |
| SR7 | Multiple samples in years $1 ; 4 ; 5 ; 8 ; 9 ; 10$ | 1 | 1 | Yes ${ }^{\text {c }}$ |
| SR8 | Multiple samples in years $1 ; 4 ; 5 ; 8 ; 9 ; 10$ | Time-varying ${ }^{\text {a }}$ | 1 | No ${ }^{\text {b }}$ |
| SR9 | Multiple samples in years $1 ; 4 ; 5 ; 8 ; 9 ; 10$ | Time-varying ${ }^{\text {a }}$ | $P_{\text {sel2 }}^{\mathrm{a}}$ | $\mathrm{No}^{\text {b }}$ |
| SR10 | Multiple samples in years $1 ; 4 ; 5 ; 8 ; 9 ; 10$ | Time-varying ${ }^{\text {a }}$ | $P_{\text {sel2 }}^{\mathrm{a}}$ | Yes ${ }^{\text {c }}$ |

${ }^{\mathrm{a}}$ See Figure 2; ${ }^{\text {b }}$ see Figure 3a; ${ }^{\text {c see Figure 3b. }}$
sampled at different age-classes (Figure 3), and a catch-curve can be fitted for each year-class. Using simulated datasets in which sampling occurred in several years (S13-S25, Table 2), this method was applied to counts-at-age for individuals pooled by year-class. Year-class 10 could only be sampled at the age-class 0 in year 10 and was therefore always excluded. For multiple samples within a year (SR6-SR10, Table 1), counts-at-age were computed from each sample separately providing multiple counts in a given ageclass, and catch-curves were fitted to the resulting set of points. In simulations where counts-at-age did not decrease monotonically, such as those including gear selectivity and the resulting underrepresentation of younger individuals, catch-curves were fitted to the declining portion of the age data. Plots of these catch-curves are given in Supplementary Figures S4-S16. It must be noted that catch curves provide an estimate of the initial number of age-0 individuals in the population. The age of individuals in observations may range from 0 to $<1$ year and these individuals may have experienced some mortality. This is in contrast to the IBA described below which estimates the initial number of individuals in the population, before mortality.

## Year-class strength estimated using the IBA

## Mathematical basis

In the model (and in the field), the fish captured are survivors of natural mortality processes. Under constant natural mortality ( $M$, year ${ }^{-1}$ ), the number of survivors $(N)$ in a given age ( $a$, year) can be expressed as a function of the initial number of larvae ( $N_{\mathrm{Tl}}$, the total number of larvae that hatched and have not yet experienced mortality) as:

$$
\begin{equation*}
N=N_{\mathrm{Tl}} \times \mathrm{e}^{-M \times(a+1)} . \tag{8}
\end{equation*}
$$

Equation (8) applies to one year-class; however, for multiple yearclasses, the year-class of each individual must be computed [Equation (7)] and Equation (8) is applied to counts of individuals

Table 2. Configurations of the model used to simulate the population dynamics.

| Simulation | Sampling regime | Mortality | Larval <br> recruitment |
| :--- | :--- | :--- | :--- |
| S1-S12 | SR1 (January to <br> December) | $M$ | Constant |
|  | SR2 | $M$ | Constant |
| S13 | SR5 | $M$ | Constant |
| S14 | SR6 | $M$ | Constant |
| S15 | SR7 | $M$ | Constant |
| S16 | SR8 | $M$ | Constant |
| S17 | SR3 | $M$ | Constant |
| S18 | SR4 | $M$ | Constant |
| S19 | SR9 | $M_{(\text {age) }}$ | Constant |
| S20 | SR2 | $M_{\text {(age) }}$ | Constant |
| S21 | SR9 | $Z_{\text {(age,t) }}$ | Constant |
| S22 | SR9 | $Z_{\text {(age,t) }}$ | Variable |
| S23 | SR9 | $Z_{\text {(age,t) }}$ | Variable |
| S24 | SR10 |  |  |
| S25 |  |  |  |

Simulations are denoted $\mathrm{S} 1-\mathrm{S} 25$, each resulting in the production of a given simulated dataset. Sampling regimes are detailed in Table 1, mortality (Figure 2 ) is either set as constant ( $M$ ), age-dependent [ $\left.M_{(\text {age })}\right]$, or the combination of age-dependent and fishing mortality $\left[Z_{(\text {age, } t)}\right]$. Larval recruitment refers to the release of larvae in the model ( $50 \%$ in April, $25 \%$ in March and May) and was either set as constant ( $30 \times 10^{6}$ individuals per year) or as variable.
pooled by year-class $\left(H_{\mathrm{Y}}\right)$ :

$$
\begin{equation*}
N_{\left(H_{Y}\right)}=N_{\mathrm{Tl}\left(H_{Y}\right)} \times \mathrm{e}^{-M \times(a+1)} . \tag{9}
\end{equation*}
$$

The actual survival rate or probability of surviving to a given age $a\left(P_{s}\right)$ is:

$$
\begin{equation*}
P_{\mathrm{s}}=\mathrm{e}^{-M \times(a+1)} \tag{10}
\end{equation*}
$$

The total number of larvae that hatched in a given year is the ratio of the number of individuals from that year-class at a given age to their probability of surviving to that age:

$$
\begin{equation*}
N_{\mathrm{T}\left(H_{\mathrm{Y}}\right)}=\frac{N_{\left(H_{\mathrm{Y}}\right)}}{P_{\mathrm{s}}} . \tag{11}
\end{equation*}
$$

The number of larvae that hatched in a given year and resulted in the survival of one individual is therefore:

$$
\begin{equation*}
N_{\mathrm{l}\left(H_{\mathrm{Y}}\right)}=\frac{1}{P_{\mathrm{S}}} . \tag{12}
\end{equation*}
$$

The total number of larvae that hatched in a given year and resulted in the survival of $N$ individuals from that year-class at a given age can be expressed as the sum of estimates for each individual:

$$
\begin{equation*}
N_{\mathrm{Tl}\left(H_{\mathrm{Y}}\right)}=\sum_{1}^{N_{\left(H_{\mathrm{Y}}\right)}} N_{\mathrm{l}\left(H_{\mathrm{Y}}\right)} . \tag{13}
\end{equation*}
$$

In this basic example, combining Equations (12) and (13) is equivalent to using Equation (11). In reality, for an individual to be observed by the survey, it has to survive, be captured, and its age must be determined. The combination of all events leading to the observation of an individual is the probability of being observed
$\left(P_{\mathrm{O}}\right)$. The number of larvae that hatched in a given year and resulted in the observation of one individual is therefore:

$$
\begin{equation*}
N_{l\left(H_{\mathrm{Y}}\right)}=\frac{1}{P_{\mathrm{O}}} \tag{14}
\end{equation*}
$$

In a single survey sample, a given year-class is observed at a given age. Combining Equations (13) and (14) provides an estimate of the total number of larvae that hatched and resulted in the observation of $N_{(H Y)}$ individuals. If an additional sample is collected later in time, the same year-class may be observed a second time (at an older age), providing a second estimate of the initial number of larvae for that same year-class. Summing estimates as shown in Equation (13) would result in an estimate that is double the actual value. It is therefore necessary to average estimates of the initial number of fish across samples, using the number of times a year-class was sampled [ $N s_{(H Y)}$, Figure 3]:

$$
\begin{equation*}
N_{\mathrm{T}\left(H_{\mathrm{Y}}\right)}=\frac{\sum_{1}^{N_{\left(H_{\mathrm{Y}}\right)}} N_{\mathrm{l}\left(H_{\mathrm{Y}}\right)}}{N_{\left(H_{\mathrm{Y}}\right)}} . \tag{15}
\end{equation*}
$$

Since year-class strength $\left(N_{\mathrm{TI}}\right)$ is computed as the mean of estimates across samples, it is also possible to compute the range of estimates (i.e. minimum and maximum) when a year-class is sampled more than once. The range of the estimate will be shown along with the mean estimate for each year-class in the results. Once the probability of observing each individual is determined, Equations (14) and (15) must be computed to provide an individual-based estimate of yearclass strength.

## The probability of being observed ( $\mathrm{P}_{\mathrm{O}}$ )

For an individual, the probability of being observed is its probability of being collected and processed by the survey. The computation of this probability will depend on the organism, the sampling protocol, and the nature of the observation. Here, for a fish sampled and consequently aged by otolithometry, $P_{\mathrm{O}}$ takes the form:

$$
\begin{equation*}
P_{\mathrm{O}}=P_{\mathrm{s}} \times P_{\mathrm{pres}} \times P_{\text {sel }} \times P_{a} \times P_{\mathrm{pop}} \tag{16}
\end{equation*}
$$

where each variable is either a probability or a proportion that ranges between 0 and 1. $P_{\mathrm{s}}$ is the probability of surviving up to the time of capture and $P_{\text {pres }}$ the probability of being present at the time and location of sampling. This latter variable is defined as availability and corresponds to the probability of being at the location and depth at which the sampling gear was deployed on a given day. $P_{\text {sel }}$ is the probability of being selected by the gear, $P_{a}$ the proportion of the fish sampled that were aged, and $P_{\text {pop }}$ the proportion of the population sampled. For a trawl survey and assuming a homogeneous distribution of the population, $P_{\text {pop }}$ is the ratio of the volume of the sample to the volume of the population. The volume of the sample is determined by the size of the opening of the net and the tow duration, whereas the volume of the population corresponds to the volume of the habitat occupied by the population. $P_{\text {sel }}$ and $P_{\mathrm{a}}$ were discussed in the Individual-based population dynamics model section; we assumed that the entire population was sampled on a given sampling date, thus $P_{\text {pres }}$ and $P_{\text {pop }}$ were set to 1 , a simplification that will be revisited in the discussion. The remaining variable, $P_{\mathrm{s}}$, depends on the cumulative mortality experienced by an individual from hatch to capture; its estimation is described in the section that follows.

It is worth noting that the factors which $P_{\mathrm{O}}$ depends on can be separated into two groups: (i) the variables that represent instantaneous processes occurring at the time and location of capture ( $P_{\text {pres }}$, $P_{\text {sel }}, P_{a}$, and $P_{\text {pop }}$ ) and (ii) the probability of surviving, which depends on the cumulative mortality experienced by an individual from the time of hatching to capture $\left(P_{\mathrm{s}}\right)$. Using an age-dependent formulation of mortality (instead of one that is length-dependent) and requiring solely the length-at-capture of each individual to estimate its probability of being selected enables the IBA approach to compute year-class strength estimates irrespective of the growth history of each fish. Consequently, the approach would be equally accurate (or inaccurate) across any growth model formulation (see Supplementary material, Part 5).

## The probability of surviving $\left(P_{s}\right)$

A major simplification associated with the traditional approach to mortality computations is that the age of individuals is computed in years, i.e. rounded to an integer. For an individual hatched in April, such rounding would be valid only if it was captured in April of a consecutive year. Fish captured in other months have therefore experienced either more or less than a full year of mortality. To avoid this bias, the age of an individual may be estimated with more accuracy, provided the species is characterized by a known and short hatching period. The method described here, which is based on a monthly resolution to match the time-step of the population model, relies on the distinction between date and time. Here, a date is composed of a year (from 1 to 10) and a month (from 1 to 12), while time is the number of months elapsed since a given starting date (Figure 4).

For a given individual observed by the survey, its age-class, year of capture, and month of capture are known, and, its year-class is the difference between its year of capture and its age-class [Equation (7)]. Knowing the first year-class found in the dataset ( $H_{\mathrm{Ymin}}$ ) and the average month in which hatching occurs ( $\overline{H_{\mathrm{M}}}$ ), the time at which the individual hatched $\left(H_{\mathrm{T}}\right)$ is:

$$
\begin{equation*}
H_{\mathrm{T}}=\left(H_{\mathrm{Y}}-H_{\mathrm{Y} \min }\right) \times 12+\overline{H_{\mathrm{M}}} . \tag{17}
\end{equation*}
$$

Similarly, the time at which this individual was captured $\left(\mathrm{Capt}_{\mathrm{T}}\right)$ is:

$$
\begin{equation*}
\mathrm{Capt}_{\mathrm{T}}=\left(\mathrm{Capt}_{\mathrm{Y}}-H_{\mathrm{Y} \min }\right) \times 12+\mathrm{Capt}_{\mathrm{M}} . \tag{18}
\end{equation*}
$$

Finally, the estimated age in months of this individual is (when


Figure 4. Schematic of the elements used to estimate the age of an individual in months. Two sampled year-class are shown (black histograms), with a mean month of hatch $\left(\overline{H_{M}}\right)$ in April. Here, an individual belonging to the 1981 year-class was captured in July ( Capt $_{M}$ ) of year 1983 (Capt ${ }_{Y}$ ), at the age-class 2 ( Age $_{C}$ ). "Time" is the number of months since the beginning of the time-series, i.e. January of the first year-class found in samples (1980).
accounting for an age 0 months):

$$
\begin{equation*}
\text { Age }_{\mathrm{M}}=\text { Capt }_{\mathrm{T}}-H_{\mathrm{T}} \tag{19}
\end{equation*}
$$

Once the age in months of an individual has been estimated, its probability of surviving may be estimated with a greater accuracy than if its age were expressed in years [Equation (10)], using the number of months, an individual has survived to 12 ths of an annual mortality rate $\left(M\right.$, year $\left.^{-1}\right)$ :

$$
\begin{equation*}
P_{\mathrm{s}}=\mathrm{e}^{-\frac{M}{12} \times\left(\mathrm{Age}_{\mathrm{M}}+1\right)} \tag{20}
\end{equation*}
$$

If mortality is age-dependent $\left[M_{(\text {age })}\right.$, year $^{-1}$ ], the probability of surviving takes a cumulative form:

$$
\begin{equation*}
P_{\mathrm{s}}=\mathrm{e}^{-\sum_{0}^{\operatorname{AgeM}_{\mathrm{M}}} \frac{M_{(a g e)}}{12}} \tag{21}
\end{equation*}
$$

Each mortality rate associated with an age in years is divided by 12 and repeated 12 times in a series of monthly values. The sum of values from the start of this series (age-0 months) to $\mathrm{Age}_{\mathrm{M}}$ is then used to compute the cumulative probability of survival.

Similarly, a time-series of monthly fishing mortalities can be computed from annual values $\left[F_{(t)}\right.$, year ${ }^{-1}$ ], ranging from time 1 (i.e. January of $H_{Y m i n}$, Figure 4) to the time of the last capture in the survey. Using the sum of monthly values from the time of hatch $\left(H_{\mathrm{T}}\right)$ to the time of capture $\left(\mathrm{Capt}_{\mathrm{T}}\right)$, the cumulative probability of surviving harvesting is:

$$
\begin{equation*}
P_{s}=\mathrm{e}^{-\sum_{H_{\mathrm{T}}}^{\text {Capt }_{\mathrm{T}}} \frac{F_{(t)}}{12}} \tag{22}
\end{equation*}
$$

Ultimately, the probability of surviving given an age-dependent mortality and a time-dependent fishing mortality is:

$$
\begin{equation*}
P_{\mathrm{s}}=\mathrm{e}^{-\left(\sum_{0}^{\mathrm{Age}_{\mathrm{M}}} M_{(\mathrm{age})} / 12+\sum_{H_{\mathrm{T}}}^{\mathrm{Capt}_{\mathrm{T}}} F_{(t)} / 12\right)} \tag{23}
\end{equation*}
$$

Equations (14-16) were used to estimate year-class strength from each simulated dataset (Table 2) by (i) knowing the proportion of fish aged in each sample, (ii) computing each individual probability of being selected, (iii) estimating each individual probability of survival, and (iv) counting the number of times each year-class was sampled ( $N s$ ). Due to the large number of simulations, some model outputs (simulations S1-S12, S18, and S19) are shown and discussed in the Supplementary material along with plots of the catch-curve fits for all simulations. Figures in the Supplementary material are denoted "S_". The software used for the population model and the estimations of year-class strength was the R statistical computing environment, version 3.1.2 (R Core Team, 2015).

## Results

## Constant mortality, multiple samples (S13-S17)

When sampling occurred regularly, for example, every June, yearclass strength estimated with the catch-curve method appropriately reflected the constant trend in the actual numbers of larvae released annually (Figure 5a). However, the values were underestimated because individuals which, on average hatched in April, were sampled in June. This offset resulted in individuals that had
experienced 3 months of mortality unaccounted for in the aggregate age-classes used by the catch-curve method.

When sampling occurred irregularly in different months across years, estimates from the catch-curve method did not reflect the constancy of the actual year-class strength (Figure 5b and c), despite the excellent fit of the catch-curves to the data (Supplementary Figure S5). This was due to the ageing inaccuracy caused by rounding ages to integers, which occurred when linear regressions were fitted to the number of individuals pooled by ageclasses. Furthermore, because sampling occurred in different months across years, the magnitude of the ageing inaccuracy was variable and resulted in variable and erroneous year-class strength estimates (see also Supplementary Figure S1).

When the IBA was applied to the entire simulated dataset (Figure 5b), the number of larvae hatched in year 8 was underestimated, due to the collection of samples in the March of year 8, during larval recruitment (see also Supplementary Figure S2). Excluding age-class 0 individuals from that sample (i.e. age-class 0 individuals from year-class 8) enabled the accurate estimation of year-class strength in year 8 , which in this case was estimated from age-class 2 individuals sampled in year 10 (Figure 5c).

Similarly, for multiple samples collected in some years, the ageing inaccuracy inherent in the catch-curve method resulted in poor estimates of year-class strength (Figure $5 \mathrm{~d}-\mathrm{i}$ ), particularly if all age-classes were not observed on some sampling dates (Figure 5 g ). For the IBA, once age-class 0 individuals were excluded from March and April samples (i.e. during larval recruitment), the approach resulted in accurate estimates (Figure 5f), even if some age-classes were missing from samples (Figure 5 g ).

Unsurprisingly, a time-varying proportion of individuals aged from samples resulted in poor estimates of year-class strength using the catch-curve method (Figure 5h and i). One of the assumptions of that method is the constancy of the proportion of the population that is sampled. In this case, the resulting estimates combine biases from the ageing inaccuracy due to irregular sampling with the inconstancy of the proportion of individuals that are aged (Figure 2d). When accounting for the proportion of fish aged in each sample, the IBA successfully estimated year-class strength (Figure 5i).

## Constant mortality, length-dependent selectivity, and irregular sampling (S20)

Results for simulations with constant mortality, length-dependent selectivity, and regular sampling $(\mathrm{S} 18, \mathrm{~S} 19)$ are shown in Part 1 of the Supplementary material (Figure S3). Under constant mortality, when time-varying proportions of individuals were sampled irregularly and according to a dome-shaped length-dependent selectivity function, the catch-curve method failed to estimate year-class strength (Figure 5j and k). Using this method resulted in a spurious trend in year-class strength, which was solely due to the sampling regime. The constancy and magnitude of year-class strength were accurately captured by the IBA, as long as the age-class 0 individuals collected in March and April were excluded from samples.

## Age-dependent mortality (S21-S22)

When survey data were collected from a population simulated under age-dependent mortality and if mortality was assumed constant in the IBA, both the catch-curve method and the IBA failed to estimate year-class strength accurately (Figure 6a). Using an age-dependent mortality rate in the population model expectedly resulted in incorrect estimates of year-class strength from the catch-curve method


Figure 5. Year-class strength. The number of larvae released in the population model annually (clear grey), the catch-curve (black) and individual-based (mean: dark grey; range: horizontal segments) estimates are overlaid. Crosses denote year-classes sampled at only one age-class. The locations of triangles symbolize dates of sampling. Stars denote cases where age-class 0 individuals were excluded from samples collected in March and/or April. For simulation S17, as part of the individual-based estimation, the proportion of individuals that were aged from samples ( $P_{\mathrm{a}}$ ) was either set to $1(\mathrm{~h})$ or to the values prescribed during the simulation (i). Details on each simulation ( $\mathrm{S}_{-}$) are given in Table 2.
(Figure 6), as a consequence of the non-linearity of the logtransformed numbers-at-age (e.g. Supplementary Figure S12). Attempting to compute the probability of survival using a constant mortality rate of 0.83 year $^{-1}$ within the IBA resulted in a parabolic trend of year-class strength (Figure 6a). This was due to the differing age composition of year-classes in samples, where, for instance, year-class 1 was represented by older individuals for which the probability of survival was underestimated (i.e. their mortality was overestimated, Figure 2a). If the age-dependent mortality rates are known, computing each individual's probability of survival
according to its age in months and the corresponding cumulative age-dependent mortality [Equation (21)] resulted in accurate yearclass strength estimates (Figure 6b), even under a complex sampling regime (Figure 6c).

## Combined age-dependent and time-varying fishing mortality (S23-S25)

Including a time-varying fishing mortality in the population model without including it in the IBA resulted in underestimates for yearclasses hatched during or before the occurrence of fishing


Figure 6. Year-class strength. The number of larvae released in the population model annually (clear grey), the catch-curve (black) and individual-based (mean: dark grey; range: horizontal segments) estimates are overlaid. Crosses denote year-classes sampled at only one age-class. The locations of triangles symbolize dates of sampling. Stars denote cases where age-class 0 individuals were excluded from samples collected in March and April. For simulation S21, as part of the individual-based estimation, the individual probability of survival ( $P_{5}$ ) was either computed with a constant mortality set at 0.83 year $^{-1}$ (a) or according to the age-dependent mortality prescribed during the simulation (b). For simulation S 23 , as part of the individual-based estimation, the individual probability of survival ( $P_{s}$ ) was either computed according to the age-dependent mortality prescribed during the simulation (d) or according to the combined age-dependent and time-varying fishing mortality prescribed during the simulation (e). Details on each simulation ( $\mathrm{S}_{-}$) are given in Table 2.
(Figures 2 b and 6 d ). The inclusion of the cumulative fishing mortality experienced by each individual from hatch to capture [Equation (23)] allowed the accurate estimation of year-class strength (Figure 6e).

Ultimately, the IBA was able to estimate year-class strength when larval recruitment varied from year to year (Figure 6f), even if some age-classes were not collected in survey samples (Figure 6g). Under such conditions, the catch-curve method produced spurious estimates and was, in particular, unable to distinguish strong from weak year-classes.

## Effect of variability in natural mortality

Adding increasing levels of variability in natural mortality during simulations did not drastically bias IBA estimates (Figure 7). Under maximum variability ( $M_{\mathrm{var}}=100 \%$ ), the root mean square error of estimates reached a maximum at around $3 \times 10^{6}$ (Figure 71) corresponding to $10 \%$ of the average annual simulated year-class strength (ca. $30 \times 10^{6}$ ). Overall, increasing levels of variability resulted in increasingly overestimated year-class strength when using the IBA (Figure 7b-k). This is due to the non-linear (saturating) relationship between the probability of dying and the instantaneous mortality rate [Equation (3)]. For example, if we consider an annual mortality rate of $M=2.0$ year $^{-1}$ and mortality variability of $M_{\mathrm{var}}=70 \%$, the monthly probability of dying [Equation (3)] will range between 0.0488 and 0.2467 , which averages to 0.1477 and is $<0.1535$, the actual monthly probability of dying under $M=$ 2.0 year $^{-1}$. The effect of this non-linearity becomes more prevalent as the variability in mortality increases.

## Discussion

## Limitations of the catch-curve method

This study shows that rounding the age of individuals into integers may lead to erroneous mortality estimates. This raises the question of how frequently this may occur. For an individual that may have hatched in any month of a given year and may have been captured in any month of the following year, there are 12 pairs of months for which the individual will have experienced mortality for an entire year (e.g. hatched in January and captured in January of the next year). Given all the possible combinations of hatching and capture months $(12 \times 12=144)$, only $8.33 \%$ of the cases in which rounding of the age of individuals into years will result in correct age designations for the purposes of mortality estimation. For the remaining $91.67 \%$ of cases, age-and therefore mortal-ity-would be either under- or overestimated. If repeated sampling does not occur at regular intervals in subsequent years, the resulting observations will comprise individuals that have experienced mortality over different durations. In this case, spurious trends in yearclass strength will result from application of the catch-curve method. This highlights the importance of the timing of sampling for mortality estimates based on counts of individuals pooled by age-classes (see also Xiao and Wang, 2007).

Apart from mortality miscalculation, pooling individuals into age-classes results in the assumption that other processes are equal among individuals, such as their probability of selection by a sampling device. Attempting to correct for the effect of gear selectivity on counts-at-age data using the average length-at-age would result in imprecise estimates due to the variability in length-at-age. All processes that act at the individual scale and determine the observation of each individual must therefore be accounted for before any form of data pooling. The precision gained by our IBA was
successfully demonstrated in all cases, including those with complex sampling regimes, variable larval supply and mortality schedules (Figure 6g).

A final issue with the catch-curve method that was brought to light by this study concerns the assumption of constancy of gear selectivity through time. As shown, the age distribution of sampled year-classes-and therefore the size distribution-depends on the timing of sampling. Consequently, the size-dependent probability of selection of a year-class may vary, and this is true even if the gear does not change over time (see also Supplementary Figure S3). For the catch-curve approach to be valid, the assumption of constant gear selectivity must therefore always be accompanied by the assumption of constancy in the size distribution of individuals.

In several cases where the catch-curve assumptions were violated, year-class strength estimates using the catch-curve method were extremely far from the true values due to the improper fit of the linear regressions to the log-transformed catch-at-age data (see Supplementary material, part 2). In such cases, the catch-at-age data were clearly not suitable for a catch-curve approach and would have been discarded in a real-world situation. These results were shown for consistency through all sampling regimes and mortality schedules tested.

## Limitations of the IBA

Although we have demonstrated that the IBA can provide more accurate estimates of year-class strength, our method has some limitations. A major limitation is the fact that all individuals are assumed to have hatched during the same month [ $\overline{H_{\mathrm{M}}}$, Equation (17)]. Although a different month could be assigned to each year class (if there was evidence to support such interannual changes), not all individuals from a given year class necessarily hatch in the same month. This approximation will affect the estimation of the time of hatching and therefore influence mortality estimates (see Supplementary Figure S17). The impact of this approximation was not apparent in our year-class strength estimates because (i) the prescribed hatching period was short and therefore the error was minimal, and (ii) the frequency distribution of hatch dates was symmetrical around April; therefore, the error made on individuals hatched in March was compensated by the error made on those hatched in May. The accurate application of the IBA for yearclass strength estimation is therefore limited to species characterized by a relatively short and symmetrical hatching period. However, this issue would be avoided if the hatch-date of individuals is known precisely; for example, hatch-date frequency distributions may be reconstructed from young fish aged with a daily resolution.

Another limitation of the IBA is that it requires quantification of the processes determining the individual probability of being observed $\left[P_{\mathrm{O}}\right.$, Equation (16)]. As a proof of concept, estimates were computed using known parameter values from a simulated population, which accounts for the accuracy of the IBA estimates we obtained. Given the form of the equations, it can be inferred that results will be particularly sensitive to mortality rates and the parameters of the selectivity function, because these parameters appear in exponential functions. These processes are difficult to quantify, but this is an unavoidable issue common to all survey data analyses, including those focused on improving catch curves (Wayte and Klaer, 2010; Thorson and Prager, 2011).

The probability of presence ( $P_{\text {pres }}$ ) and the proportion of the population sampled ( $P_{\text {pop }}$ ) were ignored in computations (both parameters were set equal to 1 ). In recent years, the understanding of environmental and life history processes driving the availability


Figure 7. Effect of variability in natural mortality ( $M_{\text {var }}$ \%; see the Simulations section) on year-class strength estimates for simulation S25 (Table 2). The number of larvae released in the population model annually (clear grey), the catch-curve (black) and individual-based (mean: dark grey; range: horizontal segments) estimates are overlaid. Crosses denote year-classes sampled at only one age-class. Results with increasing levels of variability are shown (b-k), as well as the root mean square error computed between the actual year-class strength and the individual-based estimates for each level of variability (I).
of fish to sampling gear has greatly improved (e.g. Williams and Fabrizio, 2011; Kotwicki et al., 2015). Using the method presented in this study, such processes could be included in computations at the individual level. For instance, if the ontogenetic migrations of a species were investigated through tag-recapture experiments (e.g. Dorazio et al., 1994) or through otolith chemistry analyses (e.g. Bouchard et al., 2015), the probability of presence of individuals at a sampling site could be estimated as a function of their age and the time of year. Another challenging task is the estimation of the proportion of the population that was sampled at any given time. Estimating the volume of a population and its potential changes in time is not a practical goal. An alternative is to consider this volume as constant and equal to $1 . P_{\text {pop }}$ would then be equal to the volume sampled-which may greatly vary from tow to towhence resulting in year-class strength estimates that would be a constant fraction of the actual year-class strength. The pattern of the resulting estimates of year-class strength would still be informative as indicative of relative changes in year-class strength.

Finally, the accuracy of the IBA relies on the availability of age data which limits its application to data-rich species. However, if age-at-length variability was quantified from a subset of individuals, age could be estimated and ageing uncertainty could be propagated through computations. For instance, assuming a normal distribution of age-at-length with a known mean and standard deviation for each of several length classes, such information could be used to generate multiple year-class strength estimates. The mean of estimates would then be informative of year-class strength, while their standard deviation would indicate the uncertainty arising from ageing uncertainty.

## Advantages of the IBA

Approaching the analysis of survey data with an individual focus helped to avoid the main limitation of catch-curves-and arguably other methods-namely the aggregation of individuals into integer year classes. By assuming an individual's month of hatching and knowing its month of capture, the cumulative mortality experienced by that individual can be reconstructed with greater accuracy than if yearly approximations had been used. Because each individual experiences a unique history, the pooling of individuals into ageclasses generates imprecision in mortality estimates. This is particularly true when sampling does not occur in the same month each year (in which case annual age-classes comprise a range of actual monthly ages) and when not all age-classes are collected in a given sampling event, both prevalent characteristics of field sampling. An additional advantage of the IBA lies in the computation of variables that are sample-specific such as the probability of being selected $\left(P_{\text {sel }}\right)$. As a result, the IBA may be applied to a dataset collating individuals sampled with gears characterized by different selectivities because $P_{\text {sel }}$ can be computed for each individual according to the gear with which it was sampled.

Overall, the main strength of the IBA to estimating year-class strength lies in its greater objectivity because it requires fewer assumptions about the dynamics of the sampled population. In particular, the assumption of population steady state is not required for the IBA equations to be valid. Also, in contrast to the catch-curve method, the IBA does not require subjective treatment of the catch-at-age data, such as the determination of the age of full recruitment to the fishery or the point of truncation of older ages (Smith et al., 2012). Finally, the method is straightforward to apply. A stepwise guide to applying the IBA to field survey data is provided in the Supplementary material (Part 4).

## Additional applications of the IBA

Because the processes resulting in the presence of an individual in a sample are explicitly included in computations [Equation (16)], the method enables sensitivity analyses to be carried out on each process. For example, it would be possible to investigate the effect of different selectivity functions sequentially, or, based on multiple estimates from a range of parameter values, quantify the resulting uncertainty around year-class strength estimates.

As part of our experiments, we identified an issue concerning the handling of observations of age-class 0 individuals during the hatching period. As shown (e.g. Figures 5 and 6, Supplementary Figure S2), the inclusion of such individuals in computations caused underestimates of the strength of the year-class to which they belong. These observations, however, can be used to inform the hatching period of the species sampled. For instance, if removing all age-class 0 individuals sampled in a given month from a survey dataset resulted in higher year-class strength estimates, this would indicate that the hatching period was not completed by that month.

The framework described in this study could be adapted to any organism for which age data are collected. Additional agedependent or time-varying processes thought to affect the probability of being observed $\left[P_{\mathrm{O}}\right.$, Equation (16)] could be included in computations, such as time-varying vital rates (Thorson et al., 2015; Webber and Thorson, 2015) or time-varying gear selectivity (Aires-da-silva and Maunder, 2015). The temporal resolution of computations could also be modified, for instance, to a daily resolution and enable hatch-date analyses for young fish aged by counting daily otolith growth increments.

Given the implicit inclusion in computations of the processes governing the sampling regime, the method could further inform optimal or sufficient sampling strategies. For instance, year-class strength estimates computed from a given dataset could be compared with those obtained from the same dataset after subsampling individuals at a lower frequency (e.g. every 6 months instead of every 3 months). Similarly, random subsamples of individuals within a given sampling event could be examined to determine the optimal proportion of aged fish (i.e. the minimum proportion for which year-class strength estimates remain unbiased). In addition, since the method enables a certain level of standardization of sampling effects, information provided by different surveys could be compared or combined, an otherwise difficult task (Simmonds et al., 2010).

Considering the known large fluctuations of mortality during the larval stage (Houde, 2002), the assumption of constant (and density-independent) mortality during the first year of life may be invalid (and result in spurious year-class strength estimates from both the IBA and the catch-curve method). In such cases, IBA yearclass strength estimates could be compared with information provided by egg, ichthyoplankton, and/or juvenile surveys to investigate early-stage survival and its departure from the linear approximation.

An additional application of the IBA is the estimation of recruitment to the population at any chosen age. For example, the IBA could be constrained to estimate the initial number of 12 -month-old individuals by computing the cumulative age-dependent mortality experienced between the age of 12 months and the age at capture, and, time-varying mortality experienced between 12 months after the supposed time of hatch and the time of capture [Equation (23)]. By extension, the approach could be used to estimate recruitment in each age resulting in counts-at-age corrected for the effects of the sampling regime. Consequently, approaching mortality
estimates in relative terms (Hoenig et al., 1990), ratios of corrected counts-at-age for two consecutive ages could inform the probability of surviving from one age to the other (i.e. age-dependent mortality), whereas ratios of corrected counts-at-age for the same age but over time could inform the probability of surviving from one year to the next (i.e. time-varying mortality).

## Conclusion

This study demonstrates the accuracy of a new approach to estimate year-class strength. The approach explicitly accounts for individual variability in the presence of age-dependent and time-varying mortality, and accounts for the characteristics of the sampling regime, such as sampling frequency irregularity, sample-specific processes (size-dependent selectivity, proportion of the population sampled and analysed, fish availability), and the occurrence of sampling discontinuities (random occurrence of non-observed age-classes). Overall, this study shed light on the consequences of pooling individuals into annual age-classes, a step inherent to many survey data analyses. The pooling of individuals into age classes generates imprecision by (i) misrepresenting the actual age of each individual and (ii) combining a range of individual histories. The IBA, however, requires estimates of the processes resulting in the presence of an individual in a survey dataset-including mortality-and therefore constitutes a potential additional tool for stock assessment. Comparison between year-class strength or recruitment estimates obtained through this approach and those obtained through an age-structured stock assessment could be beneficial and inform management strategies.

## Supplementary data

Supplementary material is available at the ICESJMS online version of the manuscript.

## Acknowledgements

We are grateful to two anonymous reviewers who helped to greatly improve the quality of this paper. We would like to thank J. Gregg and J. Gartland for their insights on fish ageing and otolithometry as well as C. Bonzek for sharing his knowledge on the aspects of sampling regimes in real-world surveys. Funding was provided by the Virginia Environmental Endowment. This paper is Contribution No. 3542 of the Virginia Institute of Marine Science, College of William \& Mary.

## References

Aires-da-silva, A. M., and Maunder, M. N. 2015. Dealing with timevarying composition data in fisheries stock assessment through selectivity: adding process or simplifying? Inter-American tropical tuna commission scientific advisory committee sixth meeting.
Baranov, T. I. 1918. On the question of the biological basis of fisheries. Nauch Issledov Iktiol Inst Izv I, 1: 81-128.
Beyer, J. E., and Laurence, G. C. 1980. A stochastic model of larval fish growth. Ecological Modelling, 8: 109-132.
Bouchard, C., Thorrold, S. R., and Fortier, L. 2015. Spatial segregation, dispersion and migration in early stages of polar cod Boreogadus saida revealed by otolith chemistry. Marine Biology, 162: 855-868.
Campana, S. E., and Jones, C. M. 1992. Analysis of otolith microstructure data. In Otolith Microstructure Examination and Analysis, pp. 73-100. Ed. by D. K. Stevenson, and S. E. Campana. Canadian Special Publication of Fisheries and Aquatic Sciences, 117.
Catalano, M. J., Dutterer, A. C., Pine, W. E., and Allen, M. S. 2009. Effects of variable mortality and recruitment on performance of catch-curve
residuals as indicators of fish year-class strength. North American Journal of Fisheries Management, 29: 295-305.
Cowx, I. G., and Frear, P. A. 2004. Assessment of year class strength in freshwater recreational fish populations. Fisheries Management and Ecology, 11: 117-123.
DeAngelis, D. L., Cox, D. K., and Coutant, C. C. 1980. Cannibalism and size dispersal in young-of-the-year largemouth bass: experiment and model. Ecological Modelling, 8: 133-148.
DeAngelis, D. L., and Mooij, W. M. 2005. Individual-based modeling of ecological and evolutionary processes. Annual Review of Ecology, Evolution, and Systematics, 36: 147-168.
Dorazio, R. M., Hattala, K. A., McCollough, C. B., and Skjeveland, J. E. 1994. Tag recovery estimates of migration of striped bass from spawning areas of the Chesapeake Bay. Transactions of the American Fisheries Society, 123: 950-963.
Dunn, A., Francis, R. I. C. C., and Doonan, I. J. 2002. Comparison of the Chapman-Robson and regression estimators of $Z$ from catch-curve data when non-sampling stochastic error is present. Fisheries Research, 59: 149-159.
Fiksen, Ø., and Mackenzie, B. R. 2002. Process-based models of feeding and prey selection in larval fish. Marine Ecology Progress Series, 243: 151-164.
Fortier, L., and Quiñonez-Velazquez, C. 1998. Dependence of survival on growth in larval pollock Pollachius virens and haddock Melanogrammus aeglefinus: a field study based on individual hatchdates. Marine Ecology Progress Series, 174: 1-12.
Gudmundsdottir, A., Oskarsson, G. J., and Sveinbjörnsson, S. 2007. Estimating year-class strength of Icelandic summer-spawning herring on the basis of two survey methods. ICES Journal of Marine Science, 64: 1182-1190.
Hoenig, J. M., Pepin, P., and Lawing, W. D. 1990. Estimating relative survival rate for two groups of larval fishes from field data: do older larvae survive better than young? Fishery Bulletin, 88: 485-491.
Houde, E. D. 2002. Mortality. In Fishery Science: the Unique Contribution of Early Life Stages, pp. 64-87. Ed. by L. A. Fuiman, and R. G. Werner. Blackwell Publishing, Oxford.
Koonce, J. F., Bagenal, T. B., Carline, R. F., Hokanson, K. E. F., and Nagięć, M. 1977. Factors influencing year-class strength of percids: a summary and a model of temperature effects. Journal of the Fisheries Research Board of Canada, 34: 1890-1899.
Kotwicki, S., Horne, J. K., Punt, A. E., and Ianelli, J. N. 2015. Factors affecting the availability of walleye pollock to acoustic and bottom trawl survey gear. ICES Journal of Marine Science, 72: 1425-1439.
Maceina, M. J. 1997. Simple application of using residuals from catchcurve regressions to assess year-class strength in fish. Fisheries Research, 32: 115-121.
Methot, R. D. 1983. Seasonal variation in survival of larval Northern Anchovy, Engraulis mordax, estimated from the age distribution of juveniles. Fishery Bulletin, 81: 741-750.
Nagid, E. J., Tuten, T., and Johnson, K. G. 2015. Effects of reservoir drawdowns and the expansion of hydrilla coverage on year-class strength of Largemouth Bass. North American Journal of Fisheries Management, 35: 54-61.
Ottersen, G., and Loeng, H. 2000. Covariability in early growth and yearclass strength of Barents Sea cod, haddock, and herring: the environmental link. ICES Journal of Marine Science, 57: 339-348.
Planque, B., Johannesen, E., Drevetnyak, K. V., and Nedreaas, K. H. 2012. Historical variations in the year-class strength of beaked redfish (Sebastes mentella) in the Barents Sea. ICES Journal of Marine Science, 69: 547-552.
Pope, J. G. 1972. An investigation of the accuracy of virtual population analysis using cohort analysis. ICNAF Research Bulletin, 9: 65-74.
R Core Team. 2015. R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Ricker, W. E. 1975. Computation and Interpretation of Biological Statistics of Fish Populations. Bulletin of the Fisheries Research Board of Canada, 191. 382 pp.
Robson, D. S., and Chapman, D. G. 1961. Catch curves and mortality rates. Transactions of the American Fisheries Society, 90: 181-189.
Rose, K. A., and Cowan, J. H. 1993. Individual-based model of young-of-the-year striped bass population dynamics. I. Model description and baseline simulations. Transactions of the American Fisheries Society, 122: 415-438.
Rose, K. A., Fiechter, J., Curchitser, E. N., Hedstrom, K., Bernal, M., Creekmore, S., Haynie, A., et al. 2015. Demonstration of a fullycoupled end-to-end model for small pelagic fish using sardine and anchovy in the California Current. Progress in Oceanography, 138: 348-380.
Schnute, J. T., and Haigh, R. 2007. Compositional analysis of catch curve data, with an application to Sebastes maliger. ICES Journal of Marine Science, 64: 218-233.
Serns, S. L. 1986. Cohort analysis as an indication of walleye year-class strength in Escanaba Lake, Wisconsin, 1956-1974. Transactions of the American Fisheries Society, 115: 849-852.
Simmonds, E. J., Portilla, E., Skagen, D., Beare, D., and Reid, D. G. 2010. Investigating agreement between different data sources using Bayesian state-space models: an application to estimating NE Atlantic mackerel catch and stock abundance. ICES Journal of Marine Science, 67: 1138-1153.
Smith, M. W., Then, A. Y., Wor, C., Ralph, G., Pollock, K. H., and Hoenig, J. M. 2012. Recommendations for catch-curve analysis. North American Journal of Fisheries Management, 32: 956-967.
Stevens, D. E. 1977. Striped bass (Morone saxatilis) year class strength in relation to river flow in the Sacramento-San Joaquin Estuary, California. Transactions of the American Fisheries Society, 106: 34-42.
Tetzlaff, J. C., Catalano, M. J., Allen, M. S., and Pine, W. E. 2011. Evaluation of two methods for indexing fish year-class strength: catch-curve residuals and cohort method. Fisheries Research, 109: 303-310.

Thanassekos, S., Cox, M. J., and Reid, K. 2014. Investigating the effect of recruitment variability on length-based recruitment indices for antarctic krill using an individual-based population dynamics model. PLoS ONE, 9: e114378.
Thanassekos, S., Robert, D., and Fortier, L. 2012. An individual based model of Arctic cod (Boreogadus saida) early life in Arctic polynyas: II. Length-dependent and growth-dependent mortality. Journal of Marine Systems, 93: 39-46.
Thorson, J. T., Monnahan, C. C., and Cope, J. M. 2015. The potential impact of time-variation in vital rates on fisheries management targets for marine fishes. Fisheries Research, 169: 8-17.
Thorson, J. T., and Prager, M. H. 2011. Better catch curves: incorporating age-specific natural mortality and logistic selectivity. Transactions of the American Fisheries Society, 140: 356-366.
Tuckey, T., Yochum, N., Hoenig, J., Lucy, J., and Cimino, J. 2007. Evaluating localized vs. large-scale management: the example of Tautog in Virginia. Fisheries, 32: 21-28.
Wayte, S. E., and Klaer, N. L. 2010. An effective harvest strategy using improved catch-curves. Fisheries Research, 106: 310-320.
Webber, D. N., and Thorson, J. T. 2015. Variation in growth among individuals and over time: a case study and simulation experiment involving tagged Antarctic toothfish. Fisheries Research, in press.
Williams, B. D., and Fabrizio, M. C. 2011. Detectability of estuarine fishes in a beach seine survey of tidal tributaries of lower chesapeake Bay. Transactions of the American Fisheries Society, 140: 1340-1350.
Xiao, Y. 2005. Catch equations: restoring the missing terms in the nominally generalized Baranov catch equation. Ecological Modelling, 181: 535-556.
Xiao, Y., and Wang, Y. G. 2007. A revisit to Pope's cohort analysis. Fisheries Research, 86: 153-158.
Yoklavich, M., and Bailey, K. 1990. Hatching period, growth and survival of young walleye pollock Theragra chalcogramma as determined from otolith analysis. Marine Ecology Progress Series, 64: 13-23.

