Interim Report   **IR-04-011**

Reaction Norms for Age and Size at Maturation: Study of the Long-Term Trend (1970-1998) for Georges Bank and Gulf of Maine Cod Stocks

Sébastien Barot (sebastien.barot@bondy.ird.fr)
Mikko Heino (mikko.heino@imr.no)
Loretta O'Brien (loretta.o'brien@noaa.gov)
Ulf Dieckmann (dieckmann@iiasa.ac.at)

Approved by

Leen Hordijk
Director, IIASA

February 2004

Contents

Abstract

Average age and size at maturation have decreased in many commercially exploited fish stocks during the last decades. This phenomenon could be either a direct phenotypic response to some environmental variation, or the evolutionary consequence of some selective pressure. Traditionally used maturation indices, i.e. the age and size at which 50% of individuals are mature, are not appropriate to assess the causes of changes in maturation because they are influenced, in addition to maturation per se, by growth and survival. To make up for this shortcoming, we use a reaction norm based approach to disentangle evolutionary changes and phenotypic plasticity. A method is presented to estimate the reaction norm for age and size at maturation from data commonly gathered for the management of fisheries. This method is applied to data on Georges Bank and Gulf of Maine stocks of Atlantic cod (*Gadus morhua*). The results show that maturation reaction norms in these stocks have shifted significantly downwards, resulting in a tendency to mature earlier at smaller size. These findings support the hypothesis that an evolutionary trend, probably caused by high fishing mortalities, is partially responsible for the observed decrease in age and size at maturation in these cod stocks. Two independent reasons justify this interpretation. First, there is no corresponding trend in growth that would suggest that improved feeding conditions could have facilitated maturation. Second, the results are based on maturation reaction norms, from which the known confounding effects of the growth and mortality variations are removed. Consequences of fisheries-induced evolution for the sustainability of the fishery are discussed.

Acknowledgments

About the Authors

Sébastien Barot
Adaptive Dynamics Network
International Institute for Applied Systems Analysis
2361 Laxenburg
Austria
and
IRD-LEST
32 Avenue H. Varagnat
93143 Bondy cedex
France

Mikko Heino
Institute of Marine Research
PO Box1870 Nordnes
5817 Bergen
Norway
and
Adaptive Dynamics Network
International Institute for Applied Systems Analysis
2361 Laxenburg
Austria

Loretta O'Brien
National Marine Fisheries Service
Northeast Fisheries Science Center
Woods Hole
Massachusetts 02543
USA

Ulf Dieckmann
Adaptive Dynamics Network
International Institute for Applied Systems Analysis
2361 Laxenburg
Austria

Reaction Norms for Age and Size at Maturation: Study of the Long-Term Trend (1970-1998) for Georges Bank and Gulf of Maine Cod Stocks

Sébastien Barot
Mikko Heino
Loretta O'Brien
Ulf Dieckmann

Introduction

Life-history parameters of a given species, such as age and size at first reproduction, survival rate or the number of offspring, vary in space and time (Roff 1992, Stearns 1992). They are partially genetically determined and evolve according to selective pressures. They also depend on environmental variations through phenotypic plasticity, which often has an adaptive value on its own (Stearns 1989, Scheiner 1993). Life-history parameters are directly linked to the fitness of individuals and to the dynamics of their population. Consequently, it is important to understand the relative influence of the direct and indirect environmental effects on life-history traits, i.e. the respective influence of phenotypic plasticity and evolution. Short-term experiments have often been accomplished to assess the respective influence of phenotypic plasticity and genetic differences on the variability of some life-history traits (Sorci et al. 1996, Sultan 1996, Pigliucci et al. 1997, Rohr 1997, Purchase and Brown 2001). However, few studies have analyzed a long-term trend in a life-history trait, and even fewer tried to infer the causes of such a trend, mostly because suitable data are seldom available.

We contribute to filling this gap by taking advantage of long-term data collected to support the management of commercially exploited fish stocks (Hilborn and Walters 1992). We chose to study age and size at maturation because these parameters are important for fitness (Roff 1992, Stearns 1992). For individuals, age and size at maturation influence the number of reproductive events and age-specific fertility. At the population level, they determine the size and age distributions of reproducing individuals, and influence population reproductive potential because fecundity is usually size-dependent (Roff 1992, Stearns 1992), but also often age-dependent in fishes (Trippel 1998, Trippel 1999). As a consequence, age and size at maturation influence strongly population dynamics and potential yields in stocks that are commercially fished.

Reversing an evolutionary trend requires the selective pressure to be reversed for a long period, while a phenotypic trend is reversed rapidly if environmental conditions come back to their initial state. Thus, distinguishing phenotypic plasticity and

evolutionary changes in age and size at maturation is important because evolutionary changes, when undesirable, are much more difficult to reverse than plastic changes. Taken together, we aim at answering a single question that has both theoretical and practical implications: Can age and size at maturation of a fish population evolve significantly over a period of few decades? This question is interesting per se to improve our understanding on life-history evolution. It is also important to answer this question to manage fish stocks and particularly to assess the sustainability of current management strategies.

In fisheries science, maturation is usually described by two indices: the age at which 50% of individuals are mature (A_{50}), and the length at which 50% of individuals are mature (L_{50}) (Jørgensen 1990, Chen and Paloheimo 1994, Morgan and Colbourne 1999, O'Brien 1999). In many fish stocks these indices suggest that maturation schedules have changed during the last 30 years: fish reproduce younger and younger, and at smaller and smaller sizes (Jørgensen 1990, Rijnsdorp 1993a, Rijnsdorp 1993b, Morgan et al. 1999, O'Brien 1999). These shifts in maturation could be due to purely phenotypic changes resulting from long-term trends in temperature, population density, or any other relevant aspect of the ambient environment. Population density, which influences many relevant parameters such as food availability, is actually a good candidate because it has often decreased due to high fishing pressures. This is likely to partially relax density-dependent effects that decrease growth rates and delay maturation. The second explanation for the trend in A_{50} and L_{50} is that high fishing rates could also cause genetic decreases in age and size at maturation, because fishing always alters the pattern of size- and age-dependent mortality (Rijnsdorp 1993a, Law 2000, Stokes and Law 2000, Ratner and Lande 2001). These two explanations are not mutually exclusive.

The A_{50} and L_{50} indices are usually estimated using a logistic regression to predict the probability of being mature as a function of age or size. The curves describing this probability as a function of age or size are called maturity ogives. We will therefore refer to the approach using these indices to characterize maturation changes as the ogive approach.

A_{50} and L_{50} describe the maturation process only indirectly. First, theoretical models show that maturation should depend both on age and size (Roff 1992, Stearns 1992). Empirical data support these findings (Stearns 1992, Rijnsdorp 1993b, Heino et al. 2002b). Consequently, it would be useful to combine the information given by A_{50} and L_{50} in a single object. Second, these indices describe the probability of being mature, which depends not only on maturation processes, but also on survival and growth before and after maturation. For example, a decrease in the survival rate of mature individuals would decrease the probability of being mature at age or size, even if the maturation process does not change (Heino et al. 2002b).

Characterizing the maturation process with the probability of maturing addresses the problems highlighted above. When calculated as a function of both age and size, the probability is conditioned on age and size and, therefore, allows the characterization of the maturation process independently from the processes of growth and survival. This probability corresponds to the probabilistic extension (Heino et al. 2002b) of the

classical reaction norm for age and size at maturation (Stearns 1992). We will refer to this approach, based on maturation probabilities, as the reaction norm approach.

A complete reaction norm for age and size at maturation – for short, maturation reaction norm – is constituted by the set of curves describing the probability of maturing as a function of age and size; to illustrate maturation reaction norms, it is often convenient to display only the size at which the probability of maturing is 50% against age, the so-called reaction norm midpoints. Unlike traditional reaction norms, which describe the changes in a phenotypic trait as a function of environmental variables, maturation reaction norms do not explicitly involve any environmental variable (Stearns and Koella 1986, Stearns 1989). This reaction norm interpretation assumes that environmental variability influencing maturation always results in some growth variations, and conversely, that these variations are mostly due to environmental variability. Under these assumptions, each point of the size-age space corresponds to a point of a growth trajectory characterized by a mean growth rate, which is determined by the past environmental conditions. This justifies the name of reaction norm. It is further assumed that individuals mature in a probabilistic way when their growth trajectories pass through the maturation reaction norm (Heino et al. 2002b). As for reaction norms in general, we can assume that two populations are genetically different if they have different maturation reaction norms. Estimating reaction norms for age and size at maturation is thus useful for disentangling the direct reversible effect of environmental variations (phenotypic plasticity) and possible genetic changes.

We have, in another paper (Barot et al. 2003), developed a novel method to estimate reaction norms for age and size at maturation when data on both mature and immature individuals are collected annually. The method has been validated and shown to be robust against violations to underlying simplifying assumptions by applying it to artificial datasets. Unfortunately, the method requires very large samples for the estimations to be robust, thus greatly restricting its applicability.

Our goal in this paper is twofold. First, we describe an improvement in the estimation method of Barot et al. (2003) that allows for the use of smaller sample sizes, and thereby opens ways for many concrete applications. This is made possible by combining data across several cohorts and by assuming some similarity in the shape of the reaction norm across the cohorts. Second, we demonstrate the utility of our improved estimation method by applying it to two cod (*Gadus morhua*) stocks in the Northwest Atlantic, from the Gulf of Maine and from Georges Bank. Estimation of reaction norms for age and size at maturation allows us to get better understanding of the nature of maturation changes that has been documented in these cod stocks (O'Brien 1999). We use artificial data both to test the robustness of the method and to facilitate the interpretation of the results. Lastly, we discuss the biological interpretation and the fishery implications of these results.

Material and methods

Biological and environmental data

The method used here to estimate maturation reaction norms requires that a representative sample of mature and immature individuals is collected annually, and that their age, size, and maturity status are determined (Barot et al. 2003). We examine data on two cod stocks, Georges Bank and Gulf of Maine (hereafter GB and GM), of the Northwest Atlantic (Fig. 1). The exploitation rate on both stocks has increased over the last four decades, first due to distant-water fleets (1960-1970), and subsequently due to US fisheries (Serchuk et al. 1994). The data set is obtained from bottom trawl surveys conducted each spring since 1968 by the Northeast Fisheries Science Center (Azarovitz 1981, O'Brien and Munroe 2000). About 13000 fish have been collected, corresponding to an average of 299 and 140 fish respectively sampled each year on Georges Bank and in the Gulf of Maine. In the age range of 1-5 years, annual sample sizes at age are about 60 and 30 individuals respectively from Georges Bank and the Gulf of Maine. Sampling is random but stratified by length so that comparable numbers of fish are sampled in each one-centimeter-wide length class; this ensures that a wide length range can be covered without sampling very large numbers of fish.

The most important requirement of the estimation method is that proportions of mature fish within each age and length class should be unbiased. The estimation is therefore insensitive to variations in sampling intensity or size selectivity of sampling gear that only affect numbers of observations in each age and length class but leave maturity proportions unchanged. Gear selectivity on maturity within a length class would potentially introduce a serious bias, but such an effect is not known for the cod stocks studied here. We therefore conclude that the sampling has been adequate for the purposes of this paper.

We examine two environmental variables that could influence maturation either directly or indirectly. The first variable, the spring bottom temperature anomaly, describes variations in the water temperature, which is an important factor of the physical environment of cod. This anomaly is computed as the difference between the observed temperature and a long-term average (Holzwarth and Mountain 1990) and is estimated from temperature data measured during the spring bottom trawl survey (O'Brien 1999). The second variable, the spring stratified mean weight per tow, is an index of the stock biomass. It allows assessing density-dependent effects on maturation, which could, for example, be due to a decline in food availability when the stock is abundant (O'Brien 1999, O'Brien and Munroe 2000).

General description of the estimation method

The probability of maturing at age and size, $m(age,size)$, can be calculated from estimations of the probabilities of being mature at age and size, $o(age,size)$, and from estimations of the mean annual growth at age, $\Delta size$ (Barot et al. 2003):

$$m\left(age, size\right) = \frac{o\left(age, size\right) - o\left(age - 1, size - \Delta size\right)}{1 - o\left(age - 1, size - \Delta size\right)}. \tag{1}$$
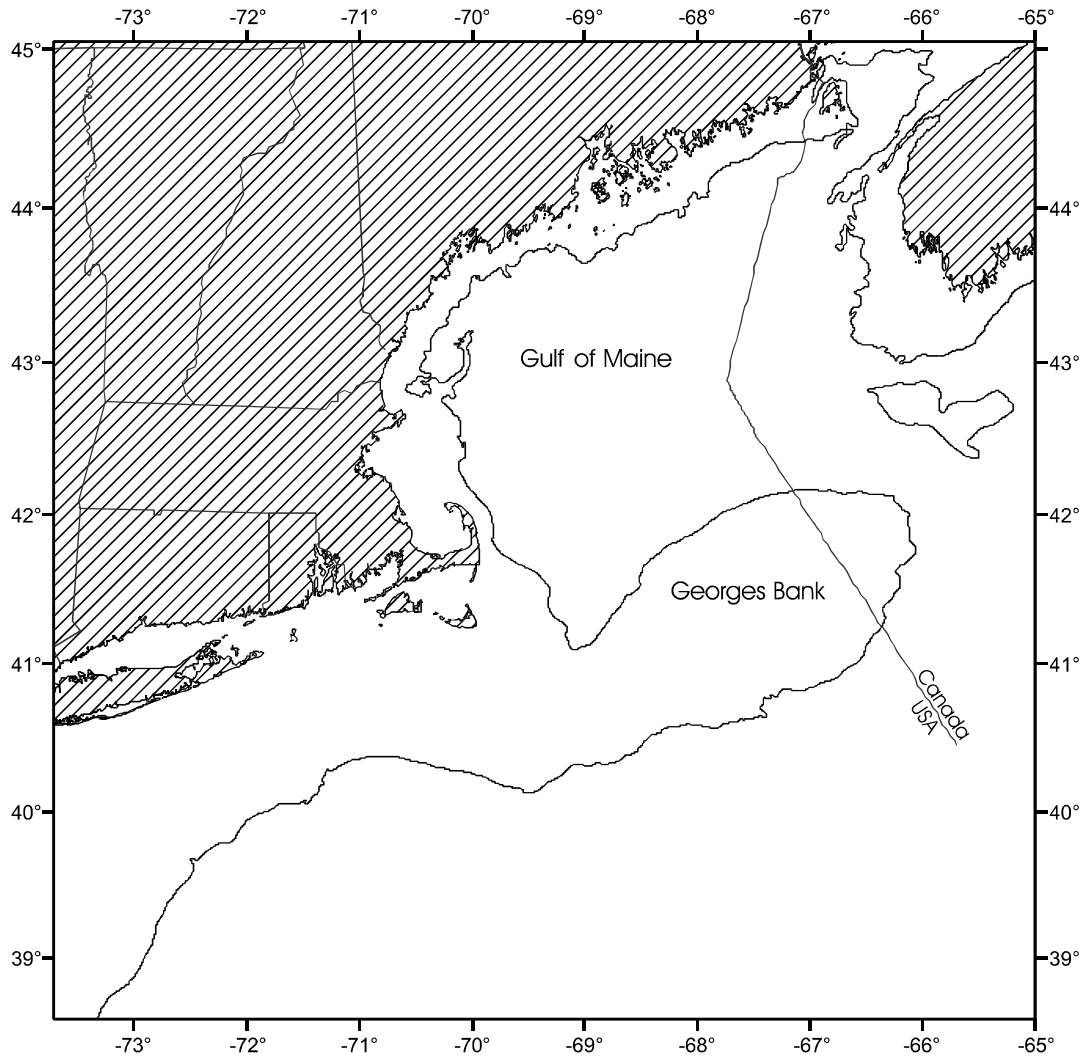
Figure 1. Geographical ranges of the two cod stocks.

This equation is strictly valid only under the assumption that immature and mature individuals have, within an age class, the same survival and growth rates. However, we have shown that the estimation is robust to violations of these assumptions (Barot et al. 2003). The full estimation method involves four steps: (1) $o(age,size)$ is estimated through a logistic regression. (2) $\Delta size$ is estimated as the difference between the mean size at age for two consecutive ages. (3) $m(age,size)$ is computed using Eq. 1. (4) An optional step is to summarize the array of probabilities $m(age,size)$ by a few parameters, e.g. the reaction norm midpoints (the sizes at which the probability of maturing is 50%); this last step is particularly useful for getting parameters that can be easily used to compare the reaction norms of different stocks or different cohorts. The first possibility is to fit a logistic regression model, and to describe the reaction norm by the model-predicted reaction norm midpoints. The second possibility is to estimate the reaction norm midpoints by interpolation between the sizes that lead to the probabilities of maturing immediately superior and inferior to 50%. Our preliminary analyses showed that the interpolation method is more robust because the logistic curve may not always fit well the estimated probability of maturing.

5

The estimation procedure outlined above has previously been applied independently to individual cohorts (Barot et al. 2003). To be able to robustly estimate reaction norm midpoints, at a minimum 100 individuals must be sampled for each considered age and cohort (Barot et al. 2003). If the probability of maturing at an age where few individuals mature (either because most of them are already mature, or conversely, because they tend to mature later) is to be estimated, even larger samples are required. To improve the performance of estimation when samples are small, we utilize information simultaneously on all available cohorts by describing the maturity ogive with a single logistic model.

Estimation of age and size-based maturity ogive

The probability of being mature at age a, and size s, for an individual of cohort c can be estimated using the following logistic regression model (Collett 1991):

$$\mathrm{logit}\left(o\left(age, size, coh\right)\right) = c_0 + c_{1,i} coh_i + c_{2,j} age_j + c_{3,i,j}\left(coh_i \times age_j\right) + c_4 size + c_{5,i}\left(coh_i \times size\right)$$
$$+ c_{5,j}\left(age_j \times size\right) + c_{6,i,j}\left(coh_i \times age_j \times size\right), \tag{2}$$

where $\mathrm{logit}(o) = \ln \dfrac{o}{1-o}$ is the logit link function. In this model cohort and age are considered as factors (discrete variables) and size as a variate (continuous variable). This statistical model is a full model: all possible interactions between the three independent variables, age, size, and cohort, are considered. This full model corresponds to the unconstrained estimation of maturity ogives for each cohort and age (c_0, $c_{1,i}$, $c_{2,j}$ … are the parameters to be estimated, i and j are respectively the indices for the cohort and the age).

Estimation of a full model should lead to the least-biased results, but it is not robust when sample sizes are low (Barot et al. 2003): when too many parameters are estimated relative to the sample size, standards errors of the estimated parameters increase and parameter estimates may become unstable. To reduce the required sample size, one must make assumptions on the common shape of the reaction norms of the different cohorts and on the effect of size across ages and cohorts. Technically, there are two solutions: reducing the number of estimated parameters (i.e. assuming that some of the constants of Eq. 2 are equal to zero), or considering age or cohort as variates.

Estimation of annual growth

Estimates of growth rates are obtained by computing the mean size at age for each cohort, and then subtracting the mean sizes of consecutive years. These values can be "smoothed", for example, using a linear model omitting the interaction between age and cohort. Our preliminary analyses showed that this was not necessary, and moreover, that reaction norm estimations are not sensitive to this choice.

Calculation of confidence intervals and randomization tests

The estimation method is based on several successive statistical analyses involving an intermediate calculation step (Eq. 1) that combines the results of the previous statistical steps. Hence direct derivation of confidence intervals or statistical tests is not possible. To surmount this problem we use bootstrap and randomization approaches (Manly 1991).

Confidence intervals are computed by bootstrapping. In a given cohort, when n_{age} individuals have been sampled at age, n_{age} individuals are chosen at random with replacement. This resampling is repeated for each cohort. The resulting resampled data set is used to derive the reaction norms of the different cohorts and their midpoints (Barot et al. 2003). The process is repeated to obtain 1000 replicates and the resulting distribution of the estimated midpoints is used to derive 95% confidence intervals (Manly 1991).

We use randomization to test statistical hypotheses on reaction norms. For example, to test whether males and females have different reaction norms for age and size at maturation, observed sex values were permutated randomly among individuals for each cohort and each age. Repeating this step for all ages and cohorts leads to a new data set for which any difference between the maturation behavior of females and males would only arise by chance. Such a data set is used to compute the probability of maturing independently for males and females. The last step (see above, step 4 in the general description of the method) is to model the probabilities of maturing, independently for each age, through a logistic regression model incorporating a sex effect and a cohort effect, both variables being considered as factors (discrete variables, k is the index for the sex), and size being always considered as a variate:

$$\text{logit}\big(m(size)\big) = c_0 + c_{1,i}coh_i + c_{2,k}sex_k + c_3 size .$$

This randomization procedure is repeated 1000 times and statistic values, here the likelihood-ratio χ^2 testing the sex effect, are collected. The same calculations are applied to the original data, without randomization. A given effect is then considered to be significant for a given age if less than 5% of randomizations leads to higher values of the test statistics than the one computed for the original data.

The same randomization procedure was used to test for a stock effect and a cohort effect using the following models as a last step:

$$\text{logit}\big(m(size)\big) = c_0 + c_{1,i}coh_i + c_{2,l}stock_l + c_3 size \text{ and}$$

$$\text{logit}\big(m(size)\big) = c_0 + c_1 size + c_2 coh .$$

In the first model both cohort and stock are considered as factors (discrete variables, l is the index for the stock). In the second model, cohort is used as a variate to test for the existence of a linear temporal trend, not merely for the existence of significant differences between cohorts. Finally, to test for the shape of the reaction norm (age effect), a randomization test based on a logistic model, taking into account age as a factor and cohort as a variate was used:

$$\text{logit}\big(m(size)\big) = c_0 + c_{1,j}age_j + c_2 size + c_{3,j}\big(age_j \times coh\big) + c_4 coh \times size .$$

Results

Growth

Growth varies from year to year but there is no long-term temporal trend (Fig. 2). This was checked for all combinations of sex and stock using a linear model including the effect of cohort (considered as a variate; F-test, P>0.05) and the effect of age. ANOVA models for both Georges Bank (GB) and Gulf of Maine (GM) stocks including the effects of age and sex did not reveal any significant difference between the size increment of males and females (F-test, P>0.05). Similarly, ANOVA models, for males and females, including the effects of stock and age, were used to test for any difference between stocks. For both sexes, yearly size increments are larger in GB than in GM (F-test, P<0.05).

Maturity ogives

Before estimating the reaction norms, a statistical model has to be chosen for the maturity ogive. Preliminary results and the study of the robustness of the estimation method (see section 'Robustness of the results' below) showed that for the GB and GM data sets the sample size at age is too low to apply the full statistical model (Eq. 2). Consequently the full statistical model has to be simplified.

   Which features of this full model must be conserved? It is a priori important to take into account the interaction between age and cohort to be able to measure the likely effects of cohort and age on the probability of being mature, but also the yearly effect of environmental variations. Preliminary analyses also showed that including cohort as a factor is necessary to be able to detect changes in the shape of the reaction norm for age and size at maturation. On the basis of these considerations, we chose to use age as a variate (continuous variable) and to include only the interaction between age and cohort:

$$\text{logit}\left(o\left(age, size, coh\right)\right) = c_0 + c_{1,i}coh_i + c_{2,i}\left(coh_i \times age\right) + c_3 size \ .$$

The consequences of this simplified model on estimations of probabilities of maturing are not straightforward. However, it is clear that it is not possible to detect changes in the inter-quartile range (defined as the width of the size interval between which the probability of maturing increases from 25 to 75%) with age or cohort, because interactions between size and age and between size and cohort were not included. However, inter-quartile ranges are not expected to be very variable as indicated by an earlier analysis using a full model for the ogive (Barot et al. 2003).

Maturation reaction norms

As an example, Fig. 3 displays the reaction norm for age and size at maturation assessed for the 1980 cohort of female GB cod. The inter-quartile range is always between 10 and 20 cm. Confidence intervals for the midpoints are narrower for ages at which most individuals mature, i.e. close to the intersection between the reaction norm and the mean

8
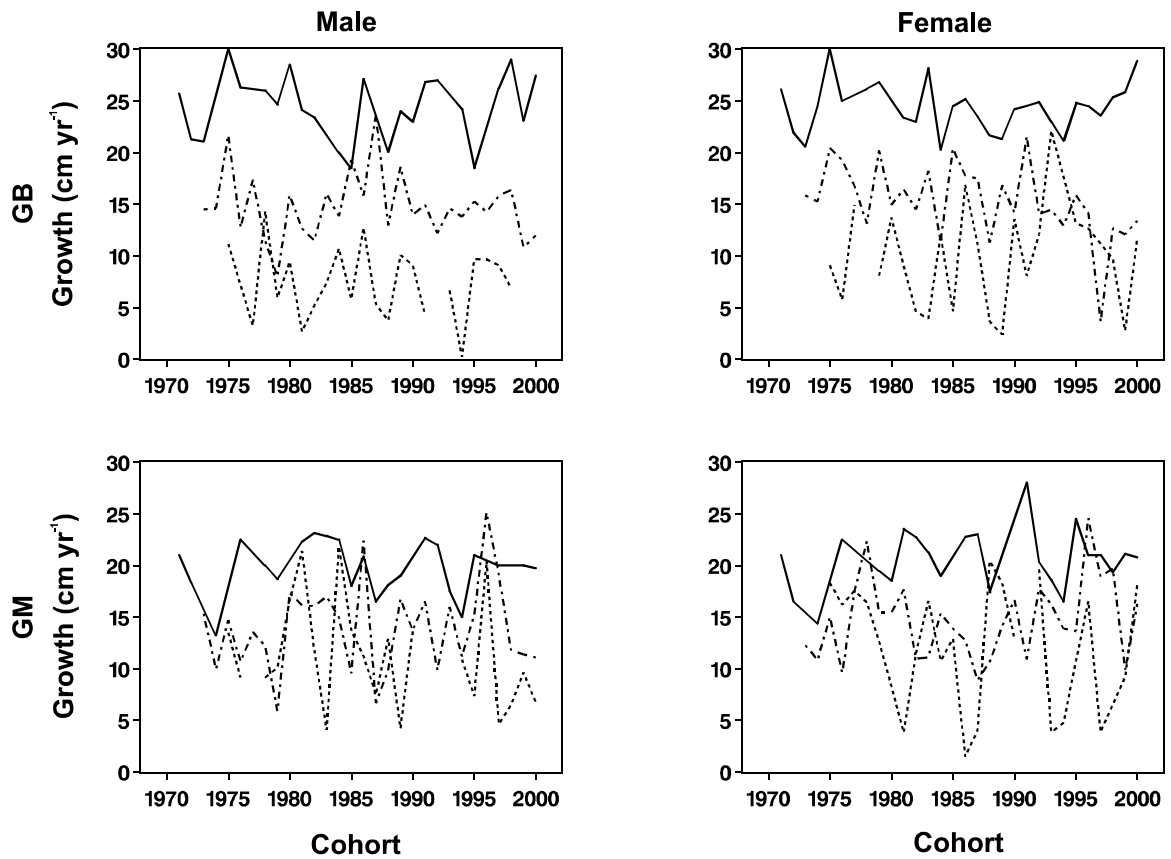
Figure 2. Annual growth increments estimated for each age and cohort. ▬ , age 1; ▬ ▪ ▬ , age 3; ▪▪▪▪ , age 5. Missing points correspond to very low growth rates that were estimated to be negative. For clarity, curves for ages 2 and 4 are not displayed, but they present the same kind of oscillations as ages 1, 3, and 5.

Figure 3. Single reaction norms (thick plain line) estimated for the females of the cohort 1980 and displayed with the inter-quartile range (thin continuous line) and bootstrapped confidence intervals for the midpoints (see method section). The mean size at age is also displayed (doted line).

9

growth curve. This simply results from the fact that more maturation data is available for these ages.

Fig. 4 displays the reaction norms estimated separately, for males and females and for the two stocks, and averaged over five years periods. There is a temporal trend towards maturat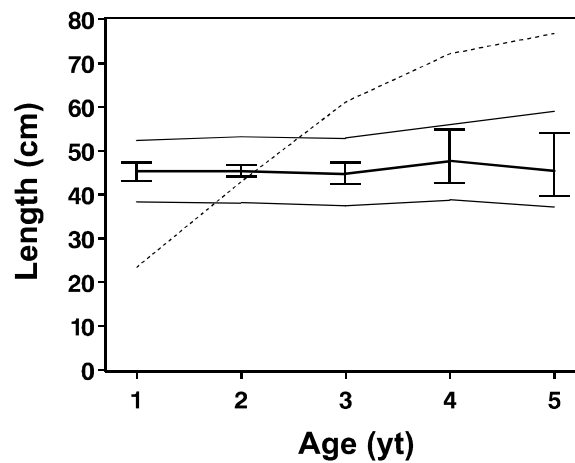ion at smaller size. This pattern is stronger for GB than for GM cod. The existence of this trend is confirmed by the results of randomization (Table 1). The trends are also visible in the midpoints for all cohorts without pooling (Fig. 5), despite the large short-term variations.

Reaction norms are horizontal in shape, or tend to be bent downwards for older ages. This pattern is significant for males and females of GB cod: old individuals have higher probabilities of maturing than young ones of the same size (randomization tests for an age effect, $P<0.05$, in these two cases predicted midpoints decrease from age 1 to 5 years). There is a trend, more evident for GB cod, towards an increase of the reaction norm slope (Figs. 4 and 5): the influence of age on probability of maturing has been increasing.

Significant differences between sexes and stocks are revealed by randomization tests (Table 1). First, the probability of maturing at age and size tends to be higher for females than for males (only two significant tests out of ten, Table 1). This means, given that the reaction norms are nearly horizontal or are slightly negatively inclined, that females tend to mature at smaller sizes and at slightly younger ages than males. Second, for age 2 the probabilities of maturing at age and size tend for both sexes to be higher for GB than for GM (Table 1), while it is the reverse for ages 4 and 5 (test significant only for females).

Environmental variations

There is a long-term trend in the biomass of the two stocks (Fig. 6, linear regression: GB, $P<0.001$, slope$=-0.39$ kg per tow yr^{-1}; GM, $P<0.001$, slope$=-0.21$ kg per tow yr^{-1}). The temperature anomaly is oscillating without a long-term trend (linear regression: $P>0.05$ for both stocks).

Robustness of the results

The estimation method has earlier been shown to be robust against violations of the simplifying assumption made to calculate the probability of maturing, that is, that growth and survival rates are similar for juvenile and mature individuals at a given size (Barot et al. 2003). Here we have made another simplifying assumption: we have estimated the reaction norms of all available cohorts at the same time, using a single, simplified model for the maturity ogive. This permits estimation of reaction norms with smaller samples. However, model simplifications might result in biases. Here we perform robustness analyses to study how simplifications of the maturity ogive manifest themselves in the estimated maturation reaction norms. We also focus on the consequences that annual variations in environment conditions and errors in determining the maturity status may have on the estimations. To do so we build artifi-

Table 1. Results of randomization tests for the effects of stock, sex and cohort (temporal trend). In the randomization approach, to test for the effect of the variable X, values of this variable are shuffled randomly among individuals that retain for the other variables their own values (see text for details). To test for an effect of cohort on maturation, cohort is used as variate (continuous variable) so that we test for a linear temporal trend in maturation. Randomization tests are applied separately for each age. + to denote that a variate as a positive effect on the probability of maturing, or M<F to denote that male have a lower probability of maturing at age and size than female. ns, non significant; *, P<0.05; **, P<0.01; GB, Georges Bank; GM, Gulf of Maine; F, female; M, male.

Effect	Stock and/or sex	Age 1	Age 2	Age 3	Age 4	Age 5
Sex	GB	ns	M<F*	ns	ns	ns
	GM	ns	ns	M<F**	ns	ns
Stock	Males	ns	GB>GM**	ns	ns	ns
	Females	ns	GB>GM**	ns	GM>GB**	GM>GB**
Temporal trend (cohort)	GB Males	ns	+**	+**	+**	+*
	GB Females	ns	+**	+**	+**	+**
	GM Males	+*	+**	+**	+*	+*
	GM Females	ns	+**	+**	+*	+*
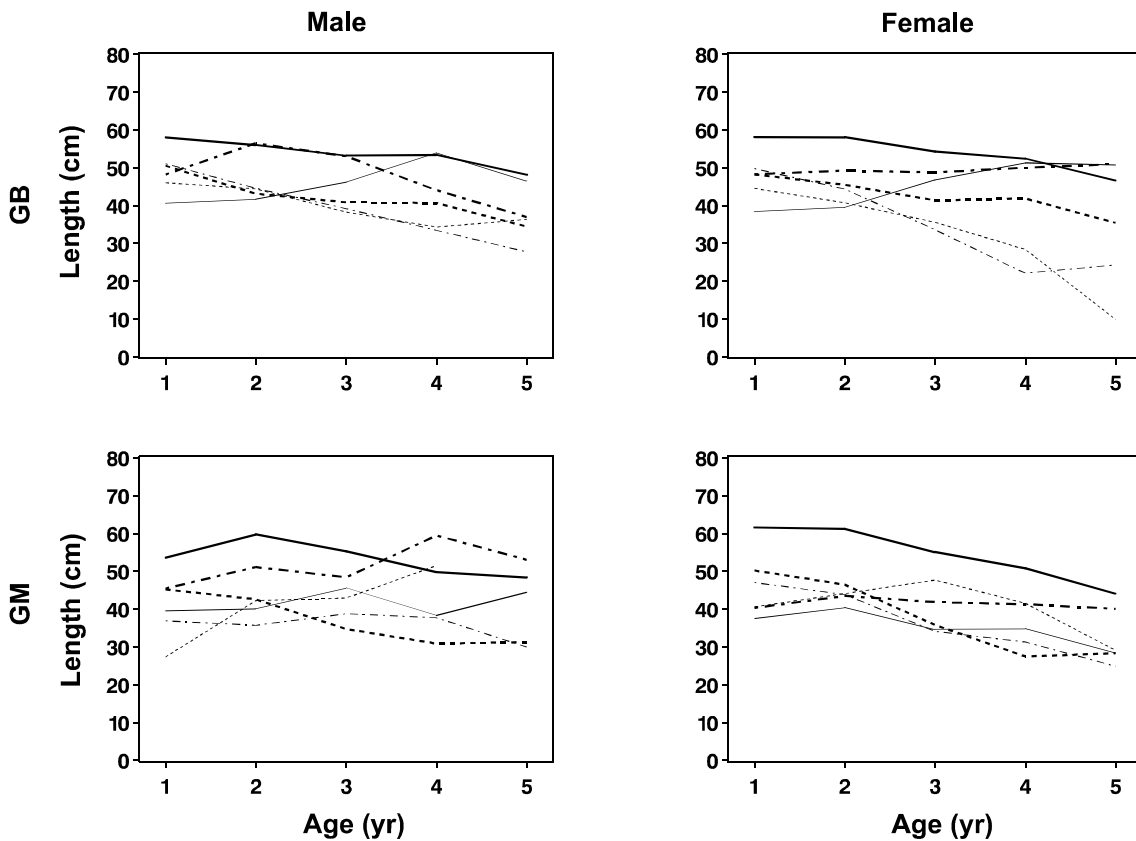
Figure 4. Reaction norms for age and size at maturation averaged over 5 years periods. Reaction norms have been estimated separately for males/females and the two stocks. ▬ , cohorts 1970-74; ▬ · ▬ , cohorts 1975-79; ▪▪▪▪ , cohorts 1980-84; ▬ , cohorts 1985-89; ▬ · ▬ , cohorts 1990-94; ······ , cohorts 1995-1999
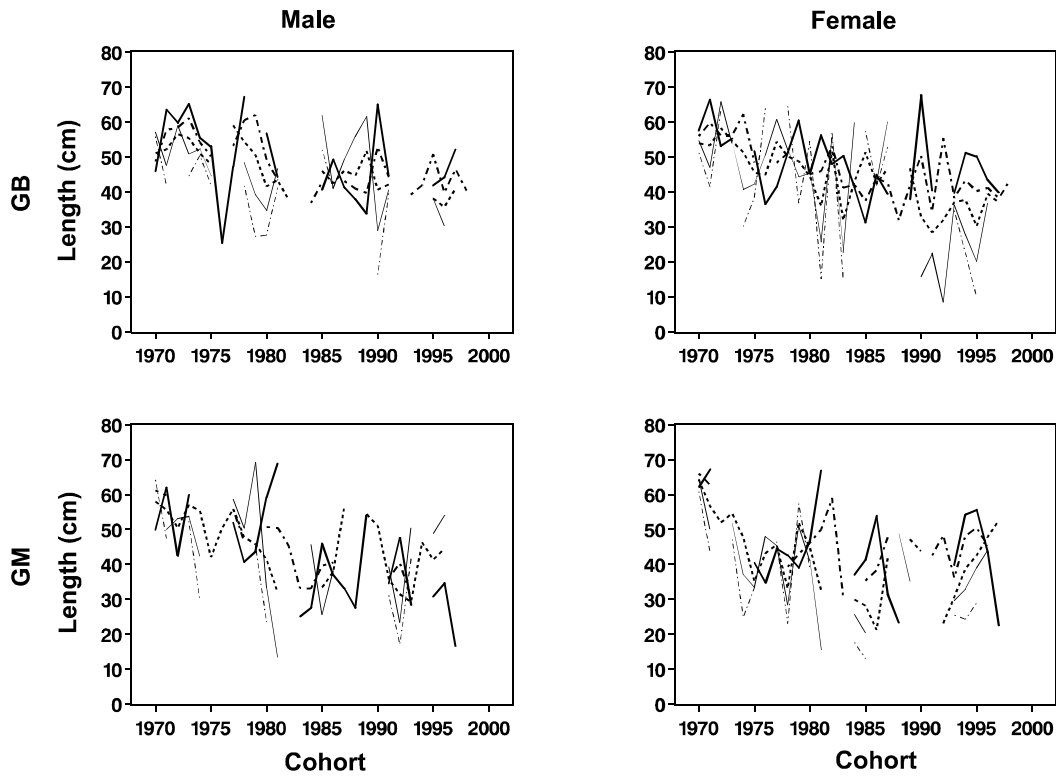
Figure 5. Temporal trend in the reaction norm for age and size at maturation midpoints. Each curve corresponds to a different age. Reaction norms have been estimated separately for males/females and the two stocks. Curves are not continuous because it was not possible to estimate some of the midpoints due to the sample size. ——, age 1; — · —, age 2; ▪▪▪▪, age 3; ——, age 4; — · — ·, age 5.
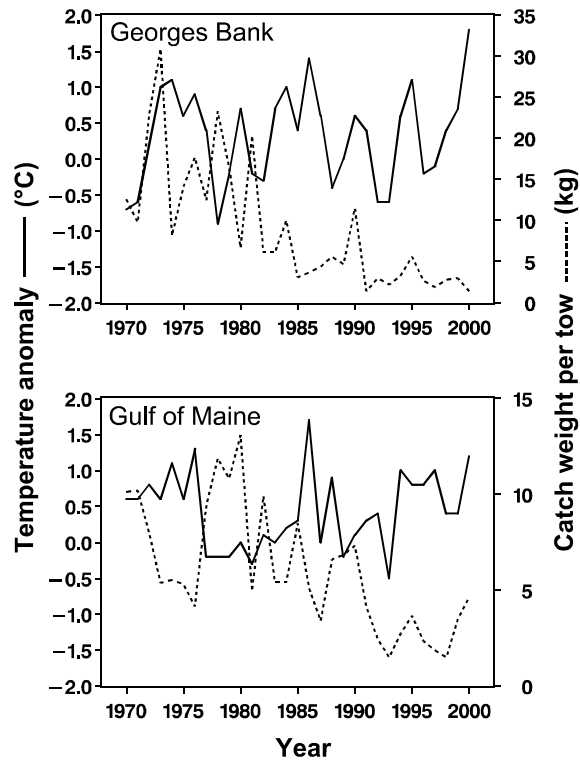
Figure 6. Temporal variations in two environmental indices: the mean weight per tow and the bottom spring water temperature anomaly. The first variable describes stock biomass variations, while the second describes variations in the climate.

12

cial data sets (Heino et al. 2002a, Barot et al. 2003) encompassing information on more than one cohort using a priori theoretical probabilistic reaction norms. These data sets are used to estimate the reaction norm using the described method and estimated and theoretical reaction norms are then compared.

Implementation

To create artificial data sets, we used the procedure described in detail by Barot et al. (2003). Data were generated allowing individuals to mature according to a given probabilistic reaction norm for age and size at maturation, to survive with a probability that may differ between juveniles and adults, and to grow deterministically. For each cohort the reaction norm is defined by the intercept of the reaction norm at the origin, its slope, and inter-quartile range (inter-quartile range of 10 cm width was always chosen, which means that an individual must grow 10 cm to increase its probability of maturing from 0.25 to 0.75). The final output is composed, for each cohort and each age, of N_{sample} randomly sampled individuals, some of which are mature and some immature. Size, age and maturity status of each individual is known. Ten replicate data sets were created for each robustness test.

The artificial data created with the procedure described above are too simplistic in one important way: the short-term environmental variability that might cause annual, cross-cohort anomalies in the tendency to mature is ignored. To implement such variability, the value of a random normal variable is added each year to the reaction norm midpoints of all cohorts. This random variable has a 0 mean, and its standard deviation (Y_{sd}) denotes the strength of the dependence of maturation on short-term environmental variability. It can be predicted that the higher is Y_{sd} the higher is the bias in the estimated reaction norms. Testing for the strength of this effect is important because the estimation method assumes that the probability of maturing depends only on age and cohort, discarding across cohort yearly effects.

At last, we checked for the robustness of the estimation method to the problem of misclassifications in maturity status. Because it can be difficult to distinguish an immature individual from a mature one for which gonads are in a resting stage, some mature individuals are probably misclassified as mature, and vice versa (O'Brien and Munroe 2000). The misclassification is likely to be "conservative": small resting fish are likely to be classified as immature, while large immature fish tend to be interpreted as resting mature fish. This probably do not significantly affect the size at which 50% of individuals are mature, but leads to an ogive with a steeper slope around that size. This was implemented using the same general procedure as for the other robustness tests and multiplying the logit of the originally estimated ogive by a factor higher than 1.

Due to the simplified statistical model used here for the ogives, the estimation method is not likely to estimate precisely the shape and the position of the reaction norms of individual cohorts. Yet, the method should estimate correctly temporal trends in the reaction norm midpoints, and simple changes in the reaction norm shape, i.e. temporal trends in the reaction norm slope. Consequently two features were used to compare theoretical reaction norms to the estimated reaction norms: the temporal trend

in the reaction norm midpoints, computed for each age by a linear regression, and the slope of the reaction norm, estimated by another linear regression. For each variable the mean and the absolute mean error were computed when possible (see below) using ten data sets.

We chose to use artificial data sets which are similar in size to the available data sets for GM cod stock: 30 cohorts, and 30 individuals sampled at age in each cohort. The following Y_{sd} values were used: 2, 4, 6, 8 and 10. Two different types of artificial data sets were created, the reaction norms being always linear. First, all reaction norms are horizontal and shift vertically for the successive cohorts (for the first cohort the midpoints values are 55 cm for all ages, while for the 30th cohort the midpoints values are all 40 cm). Second, the reaction norm of the first cohort is horizontal, but the reaction norm slope decreases gradually until it reaches a slope of -3 cm yr^{-1} in the last cohort (for the first cohort the midpoints values are 55 cm for all ages, while for the 30th cohort the midpoint at age 1 is still 55 cm, while the midpoint at age 5 is at 40 cm).

Results of the robustness tests

The first case is a vertical temporal shift of the reaction norm. This trend is recovered by the estimation method (Table 2). Even when the year effect (Y_{sd}) increases, the mean estimated temporal trend is not biased systematically although the mean error increases. Only at age 1 yr, at which only few individuals mature, increasing the year effect introduces a small bias. Mean errors in the estimated midpoints are homogeneous across the five ages. Estimations of the slope of the reaction norm (relationship between midpoint and age) are slightly biased towards negative values (Fig. 7) for all but for the higher intensity of the year effect ($Y_{sdt}=10$) for which the bias is positive. Errors in slope increase with this year effect.

In the second case, we assume the reaction norms of the successive cohorts to be more and more tilted clockwise. The temporal trend of the reaction norm midpoints is correctly detected (Table 3) when the year effect is not too high ($Y_{sd}<6$). When the year effect increases the temporal trend in the reaction norm midpoints at age is qualitatively well estimated (a decrease in the reaction midpoints is detected in all cases but one, see also Fig. 7) but errors in the reaction norm slope (relation between midpoints and age) increase: the shape of the reaction norm is less and less well estimated.

Estimations of both the temporal trend in the reaction norm midpoints and the slope of the reaction norm were found to be very robust to errors in the determination of the maturity status (Table 4). In fact, the mean estimation errors even decrease slightly when the ogive bias and the percentage of misclassification increase.

Taken together, the estimation method is robust to the violation of the assumption that there is no year effect across cohorts. The method is also robust to the misclassification of individuals into the mature and immature groups, which is the main problem likely to decrease the quality of data. In particular, our robustness tests show that the sample sizes and number of cohorts available for GB and GM cod stocks are high enough for our results to be reliable, even if the year effect on maturation is strong.

Table 2. Robustness assessment in the case of a vertical temporal shift of horizontal reaction norms (see text for details). Five cases, corresponding to an increase in the year effect (Y_{sd}) on maturation, are studied. For each age the mean temporal trend (slope of the relation between midpoints and the cohort number averaged across ten replicate data sets) is displayed as well as the mean absolute error in this trend. The mean (across all cohorts and replicate data sets) reaction norm slope (slope of the relation between midpoints and age) and the corresponding mean absolute errors are also displayed.

| | Temporal trend | | | | | | | | | | Slope | |
| | Age 1 | | Age 2 | | Age 3 | | Age 4 | | Age 5 | | | |
	Mean	Error	Mean	Error	Mean	Error	Mean	Error	Mean	Error	Mean	Error
Actual reaction norm												
	−0.50		−0.50		−0.50		−0.50		−0.50		0.00	
Estimated reaction norm												
$Y_{sd}=2$	−0.51	0.13	−0.55	0.07	−0.53	0.08	−0.60	0.14	−0.61	0.20	−0.25	0.85
$Y_{sd}=4$	−0.42	0.11	−0.48	0.10	−0.53	0.06	−0.52	0.06	−0.53	0.11	−0.30	0.83
$Y_{sd}=6$	−0.39	0.15	−0.48	0.14	−0.53	0.12	−0.46	0.09	−0.43	0.19	−0.32	1.38
$Y_{sd}=8$	−0.45	0.29	−0.66	0.27	−0.62	0.20	−0.60	0.18	−0.56	0.18	−0.77	1.60
$Y_{sd}=10$	−0.28	0.35	−0.39	0.29	−0.42	0.19	−0.36	0.23	−0.50	0.35	0.25	1.39

Table 3. Robustness assessment in the case of reaction norms of successive cohorts that are more and more tilted clockwise (see text for details). Five cases, corresponding to an increase in the year effect (Y_{sd}) on maturation, are studied. The same statistics as in Table 2 are given but here no mean value is displayed for the slope of the reaction norm (slope of the relation between midpoints and age) because there is a different slope for each cohort.

| | Temporal trend | | | | | | | | | | Slope |
| | Age 1 | | Age 2 | | Age 3 | | Age 4 | | Age 5 | | |
	Mean	Error	Mean	Error	Mean	Error	Mean	Error	Mean	Error	Error
Actual reaction norm											
	−0.10		−0.20		−0.30		−0.40		−0.50		
Estimated reaction norm											
$Y_{sd}=2$	−0.05	0.15	−0.14	0.12	−0.33	0.11	−0.36	0.11	−0.67	0.21	1.17
$Y_{sd}=4$	−0.02	0.20	−0.17	0.12	−0.31	0.10	−0.38	0.13	−0.52	0.28	1.52
$Y_{sd}=6$	−0.07	0.21	−0.15	0.11	−0.36	0.13	−0.34	0.20	−0.67	0.32	1.34
$Y_{sd}=8$	−0.12	0.18	−0.19	0.19	−0.44	0.19	−0.61	0.33	−0.47	0.35	1.73
$Y_{sd}=10$	−0.01	0.23	−0.18	0.25	−0.27	0.24	−0.43	0.30	−0.55	0.31	2.13

Table 4. Robustness of the estimation method to the misclassification of individuals into the immature and mature groups. As for Table 2, reaction norms are horizontal and shift vertically. The results are presented for a given intensity of the yearly effect on maturation ($Y_{sd}=4$, see Table 2). Four cases, corresponding to an increasing bias in the maturity ogive estimation are studied. This bias is implemented by multiplying the ogive logit by an increasing factor F, which increases the ogive steepness: multiplying the logit by a factor equal to 1.25, 1.5 and 1.75 respectively decreases a probability of being mature of 25% to 20%, 16% and 13%. In each case, ten replicate data sets have been constructed. The same statistics as in Table 2 are displayed.

| | Temporal trend | | | | | | | | | | Slope | |
| | Age 1 | | Age 2 | | Age 3 | | Age 4 | | Age 5 | | | |
	Mean	Error	Mean	Error	Mean	Error	Mean	Error	Mean	Error	Mean	Error
Actual reaction norm												
	−0.50		−0.50		−0.50		−0.50		−0.50		0.00	
Estimated reaction norm												
$F=1$	−0.41	0.08	−0.47	0.12	−0.53	0.07	−0.51	0.08	−0.53	0.15	−0.58	1.18
$F=1.25$	−0.42	0.11	−0.51	0.10	−0.51	0.06	−0.53	0.10	−0.56	0.20	−0.45	0.83
$F=1.5$	−0.42	0.11	−0.49	0.09	−0.50	0.05	−0.52	0.06	−0.53	0.10	−0.50	0.84
$F=1.75$	−0.42	0.11	−0.49	0.09	−0.51	0.04	−0.53	0.06	−0.53	0.10	−0.52	0.84
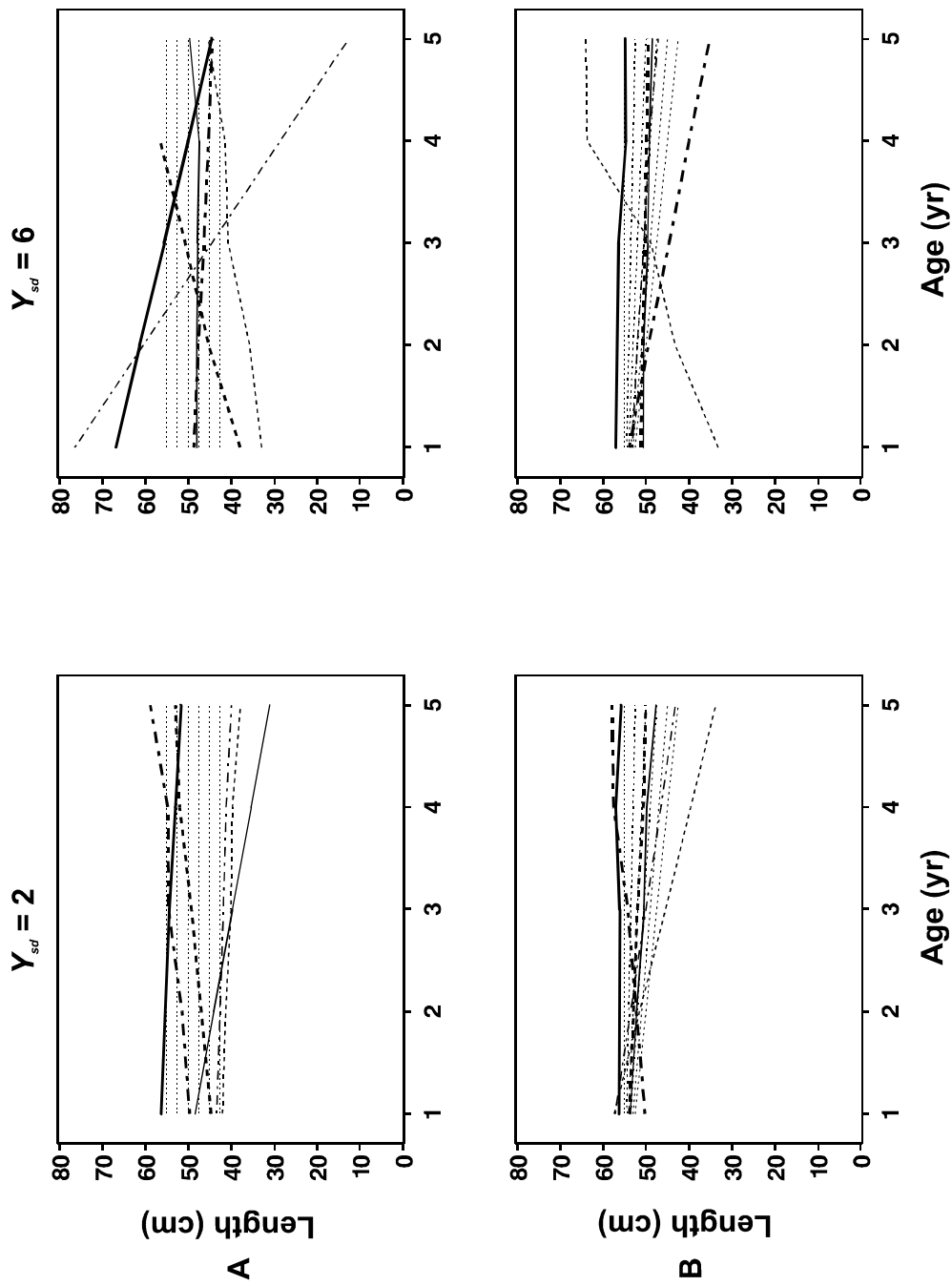
Figure 7. Graphic example of robustness tests. (A) For the first test (no age effect), theoretical reaction norms are horizontal and shift vertically. (B) For the second test (age effect), reaction norms of the successive cohort are more and more tilted clockwise. Five cases, corresponding to an increase in the year effect (Y_{sd}) on the reaction norm midpoints have been considered, and two cases are displayed here. Because the reaction norms of 30 cohorts cannot be displayed on the same figure, the reaction norms of 6 evenly spaced cohorts are displayed. Thin dotted lines, theoretical reaction norms ———, cohort 0; — · —, cohort 5; — ▪ —, cohort 10; ———, cohort 15; — · — , cohort 20; ······, cohort 25.

16

Discussion

Our first main result is methodological: we can conclude that, with data on both immature and mature individuals, presence of long-term trends in maturation reaction norms can be assessed even when annual samples are relatively small. Consequently, the methodology is now in place for utilizing the maturation reaction norm approach for many fishery data sets. This allows analyzing the long-term trend in a life-history trait, which has seldom been achieved before. The second main result is that our analyses reveal a shift of the maturation reaction norm towards lower ages and sizes at maturation in the case of GM and GB cod stocks. We discuss below the interpretation of such a trend, and emphasize that maturation reaction norms can help to better understand changes in maturation that were previously suggested by the maturity ogive approach.

Maturity ogives vs. maturation reaction norms

It must first be emphasized that the reaction norm and the maturity ogive approaches bring forward qualitatively different information. A maturation reaction norm describes the tendency to mature and is thereby only focused on the maturation process itself. A maturity ogive describes the maturity status of a stock but a maturity ogive is influenced not by one but by three processes, maturation, growth and survival. Consequently, it is easier to interpret variations in maturation reaction norms than in maturity ogives.

Variations in maturation of GB and GM cod have earlier been studied using estimations of the probability of being mature, i.e. the maturity ogives (O'Brien 1999). We have examined the same issue using the reaction norm approach, elaborating the earlier findings in some respects. First, our results reveal that the long-term trend towards lower ages and sizes of reproducing individuals also reflects a change in the maturation process per se. Second, they show that the tendency for GB cod to mature earlier than GM cod is largely caused by differences in the environment, but that there is also some differences in the maturation process. Third, in line with the earlier observation of slightly higher proportion of mature individuals at age for female than for male cod, our results indicate a somewhat higher tendency for females to mature at a given age and size than males.

Interpretation of maturation reaction norms

How should be interpreted the maturation difference between males and females? Because males and females of a given stock experience the same environment, the differences in probability of maturing must be due to intrinsic, genetic differences between males and females. Sex-specific maturation schedules could have evolved due to a difference in the reproductive energy expenditure of males and females, or in the dependence of fecundity and survival on size (Roff 1992, Stearns 1992). Our results might look surprising because it is acknowledged that males have often evolved smaller sizes and ages at maturation than females, presumably because female reproduction

requires more energy than male reproduction (Stearns 1992). Nevertheless, the observed difference between sexes is very small and only significant in two cases (Fig. 4 and Fig. 5). It would be informative to apply the reaction norm approach to other cod stocks to check if the differences in the maturity ogives between sexes (Beacham 1983, Trippel et al. 1997, Ajiad et al. 1999) correspond to maturation differences, and not to survival or growth differences.

Interpreting the temporal trend in maturation reaction norms and the differences between the two stocks is more complicated than interpreting the sex difference in maturation because environmental conditions vary in both time (Fig. 6) and space (O'Brien 1999): males and females of a given cohort experience the same environment, while different cohorts and different stocks experience potentially very different biotic and abiotic conditions.

How should be interpreted the temporal trend in maturation reaction norms? Two non-exclusive hypotheses can explain this trend. A first hypothesis is that a selective pressure has caused evolution towards low age and size at maturation. A second hypothesis is that the temporal trend in maturation is due to phenotypic plasticity. The maturation trend would then be explained by a parallel long-term trend in the environment. Stock biomass and water temperature are a priori candidate environmental variables.

Water temperature as an environmental variable explaining the maturation trend can be ruled out because it does not show any long-term trend (Fig. 6). On the contrary, stock biomass displays a decreasing trend (Fig. 6). Moreover, stock biomass is a relevant variable because it may influence cod life-history through density-dependent processes. For example, food availability may increase when biomass is low. However, the trend in the stock biomass is unlikely to be here the cause of the trend in maturation. Growth is sensitive to the same environmental variables that influence maturation (Wootton 1998). Consequently, if the temporal trend in maturation was due to phenotypic plasticity, we should also observe a temporal trend in the growth rate, but no such trend could be demonstrated here. Nor are we aware of any evidence pointing to the existence of an other environmental variable presenting a long-term trend. For these reasons, the long-term trend in maturation reaction norms and the corresponding decline in age and size at maturation has probably a significant genetic component.

How should be interpreted the maturation differences between the two stocks? The ogive approach shows clearly that cod matures earlier and at smaller sizes in GB than in GM (O'Brien 1999). However, the reaction norm approach indicates a significant difference only in four tests out of ten (Table 1). This means that the difference in maturation propensity between the two stocks is actually less important than indicated earlier, and that this difference might actually be the reverse for older ages. Thus, earlier maturation in GB than in GM cod must be caused largely by differences in the environment. GM and GB cod stocks experience different environmental conditions despite the geographic proximity (O'Brien 1999): GB constitutes a highly productive shoal averaging 50 m in depth, while GM is a deeper area with an average depth of 150 m. Moreover, the autumn water temperature is higher for GB than for GM. These differences correspond to relevant aspects of the physical environment of a cod because

they result in higher growth rates in GB than in GM (Fig. 2). It is thus possible that these environmental differences lead to plastic changes in age and size at maturation. Yet, common environment experiments suggest the existence of genetically determined differences in some growth-linked parameters between GB and GM cod stocks (Purchase and Brown 2001). Moreover, molecular studies have proved that geographically very close cod stocks can be genetically different (Ruzzante et al. 1995). The detected differences in the reaction norms, albeit small, are in accordance with these results as they indicate the presence of a small genetic difference between the stocks. The life-history trait we are studying, the maturation reaction norm, is not selectively neutral. Thus, although the two stocks are probably not genetically isolated because of movements of adult cod (Hunt et al. 1998), selection on maturation reaction norms in the two stocks could easily maintain some genetic differences between them.

We must add that the present study is one of the first ones utilizing new ideas on how maturation reaction norms should be estimated (Heino et al. 2002b), and applying these ideas to commercially fished stocks (Grift et al. 2003). The weaknesses of this approach are not yet properly understood. It is clear that not all environmental effects can be captured when expressing maturation tendency as a function of age and length (other explanatory variables, if measured, could easily be included). Nevertheless, the method removes the known effects of variations in growth and mortality from the description of maturation process. Even though residual environmental effects remain, trends in maturation reaction norms provide much stronger support for genetic changes in maturation than trends in maturity ogives that are strongly influenced by growth and mortality variations. In this sense, our method makes better use of the data. We consider that our approach is an important step forward in understanding maturation dynamics and detecting evolutionary changes in commercially fished stocks. We would like to encourage experiments with and applications to new populations (also other than fish) in order to gain further confidence that a trend in maturation reaction norms points at a genetic trend.

Can fishing pressure lead to rapid evolution?

Our analyses support the hypothesis that age and size at maturation in GB and GM cod stocks have changed genetically in response to fishing. The possibility of fisheries-induced evolution has been widely recognized since Borisov's (1978) and Ricker's (1981) pioneering work that evolution of a life-history trait could be due to the selective pressure exerted by fishing (Law and Grey 1989, Ylikarjula et al. 1999, Hutchings 2000b, Law 2000, Stokes and Law 2000, Ratner and Lande 2001). However, these ideas have mostly been developed using verbal arguments and theoretical models, with no serious attempts of empirical verification. This is the first time that some data lead so close to the demonstration that fishing-induced selection has caused an evolutionary change in a life-history trait in only three decades. More generally, it is one of the few times that the evolution of a life-history parameter at the scale of several decades is shown with field data. Earlier demonstrations of rapid evolution have mostly relied on an experimental approach (Reznick et al. 1990, Reznick et al. 1997, Thompson 1998).

Because fishing increases mortality and fishing gears are always size-selective, it is generally accepted that fishing represents a selective pressure for life-histories (Law 2000, Stokes and Law 2000, Conover, 2002, Heino and Godø 2002). It is more difficult to predict the outcome of such a selective pressure. What should be the direction of evolution for a given exploitation regime? How quick could be the consecutive evolution? Both verbal arguments and formal modeling predict that harvesting of both immature and mature fish (which is the case for Georges Bank and Gulf of Maine cod stocks) selects for low ages and small sizes at maturation (Borisov 1978, Law and Grey 1989, Ylikarjula et al. 1999, Law 2000). However, our understanding on the expected rates of changes is poor. The observed changes in the reaction norm midpoints, about 20 cm for all ages (Fig. 4 and Fig. 5), might appear too large to be due to 30 years of selective pressure on a species for which the average age at maturation is about 3 yrs. Yet, fishing mortality has been very high for Gulf of Maine and Georges Bank cod stocks (O'Brien 1999, O'Brien and Munroe 2000), resulting in selection differentials that have probably been large.

High evolutionary rates do not only require large selection differentials, but they also require high heritabilities for the studied life-history traits. Because life-history traits are directly linked to fitness, their heritabilities are often assumed to be lower than the heritabilities of morphological traits (Mousseau and Roff 1987). However, relatively high heritabilities (mean of eight fish breeding experiments: $h^2 = 0.31$) have been found for age at maturation (Law 2000), and such values should not preclude quick response to selection. Tank experiments on size-selective harvesting of daphnids and fish have also demonstrated rapid evolutionary responses (Edley and Law 1988, Conover and Munch 2002). It must nevertheless be marked that assessing the heritability of maturation parameters in the wild is the prerequisite to estimate safely how quickly can evolution proceed under fishing pressure. Modeling will then be the necessary tool to check whether the observed decline in age and size at maturation are compatible with assessed heritabilities and fishing selectivity (Ratner and Lande 2001).

Consequences for the sustainability of fisheries

Changes in maturation propensity are fundamental for understanding the long-term dynamics of commercially fished species. Many fish stocks have collapsed, probably due to the conjunction of multiple factors that generally involve overfishing (Myers and Cadigan 1995, Myers et al. 1996). This has resulted in many attempts to estimate more precisely the parameters determining stock dynamics, including their variability. These collapses have also resulted in more cautionary harvesting regimes that aim at rebuilding the stocks. In this context, the reason why some stocks, in particular the Northwest Atlantic cod stocks, are not rebuilding is unclear (Hutchings 2000a).

One explanation involves the maturation of individuals at smaller and smaller sizes, and at younger and younger ages (Hutchings 1999, Trippel 1999, Murawski et al. 2001). Such a trend is likely to reduce indirectly the reproductive potential of stocks, leading to low recruitment rates. This can be due to two mechanisms. First, old individuals spawn for a longer period than young ones, which increases the chances of larval emergence

during a peak of zooplankton abundance (Hutchings and Myers 1993). Second, smaller sizes lead to lower fecundity, and younger ages at maturation seem to lead to lower egg quality (Trippel 1998, Trippel 1999).

This scenario, linking low reproductive potentials and low age and size at maturation, constitutes another reason to assess whether age and size at maturation have changed either through phenotypic plasticity or because of a change in the genetic composition of the stock. The latter option would mean that age and size at maturation are unlikely to notably increase within a short period of time because evolution towards delayed maturation at larger size can only take place very slowly (Law and Grey 1989, Heino 1998, Law 2000). Strong selection for increased age and size at maturation would require the use of new fishing strategies designed to reverse the original selective pressure. These are probably difficult to undertake in practice. Recruitment would thus be unlikely to increase quickly after the fishery closure, and fish stocks have low chances to rebuild. Our results support this pessimistic scenario.

A decrease in age and size at maturation has been detected in many fish stocks using the maturity ogive approach, e.g., American plaice (Morgan and Colbourne 1999), North sea plaice (Rijnsdorp 1989, Rijnsdorp 1993a) and cod (Jørgensen 1990, Trippel et al. 1997). Together with our findings, these results suggest that designing sustainable fishing strategies also requires taking into account the evolutionary consequences of fishing pressure that constitute long-term feedback loops. Hence, we advise the use of management strategies that take into account the evolutionary effect of fishing, i.e. Darwinian fishing strategies (Law and Grey 1989, Heino 1998, Conover 2000, Law 2000, Stokes and Law 2000).

References

Ajiad, A., T. Jakobsen, and O. Nakken. 1999. Sexual differences in maturation of Northeast Arctic Cod. Journal of Northwest Atlantic Fishery Science **25**:1-15.

Azarovitz, T. R. 1981. A brief historical review of the Woods Hole Laboratory trawl survey time series. Canadian Special Publication in Fisheries and Aquatic Sciences **58**:62-67.

Barot, S., M. Heino, L. O'Brien, and U. Dieckmann. 2003. Estimation of reaction norm for age and size at maturity with missing first-time spawner data. IIASA, Laxenburg, Interim Report. IR-03-043

Beacham, T. D. 1983. Variability in median size and age at sexual maturity of Atlantic cod, *Gadus morhua*, on the Scotian shelf in the Northwest Atlantic Ocean. Fishery Bulletin **81**:303-321.

Borisov, V. M. 1978. The selective effect of fishing on the population structure of species with long life cycle. Journal of Ichthyology **18**:896-904.

Chen, Y., and J. E. Paloheimo. 1994. Estimating fish length and age at 50% maturity using a logistic type model. Aquatic Science **56**:206-219.

Collett, D. 1991. Modelling binary data. Chapman & Hall, London.

Conover, D. O. 2000. Darwinian fishery science. Marine Ecology Progress Series **208**:299-313.

Conover, D. O., and S. B. Munch. 2002. Sustaining fisheries yields over evolutionary time scales. Science **297**:94-96.

Edley, M. T., and R. Law. 1988. Evolution of life histories and yields in experimental population of *Daphnia magna*. Biological Jounral of the Linnean society **34**:309-326.

Grift, R. E., A. D. Rijnsdorp, S. Barot, M. Heino, and U. Dieckmann. 2003. Trends in reaction norms for maturation in North Sea plaice. Marine Ecology Progress Series **257**:247-257.

Heino, M. 1998. Management of evolving fish stocks. Canadian Journal of Fisheries and Aquatic Sciences **58**:1971-1982.

Heino, M., U. Dieckmann, and O. R. Godø. 2002a. Estimation of reaction norms for age and size at maturation with reconstructed immature size distributions: a new technique illustrated by application to Northeast Arctic cod. ICES Journal of Marine Science **59**:562-575.

Heino, M., U. Dieckmann, and O. R. Godø. 2002b. Measuring probabilistic reaction norms for age and size at maturity. Evolution **56**:669-678.

Heino, M., and O. R. Godø. 2002. Fisheries-induced selection pressures in thecontext of sustainable fisheries. Bulletin of Marine Science **70**:639-656.

Hilborn, R., and C. J. Walters. 1992. Quantitative fisheries stock assessment. Chapman & Hall, New York.

Holzwarth, T., and D. Mountain. 1990. Surface and bottom temperature distributions from the Northeast Fisheries Center spring and fall bottom trawl survey program, 1963-1987. NOAA/National Marine Fisheries Service, Lab. Ref. Doc. No. 92-08. 62 p.

Hunt, J. J., W. T. Stobo, and F. Almeida. 1998. Movement of Atlantic cod, *Gadus morhua*, tagged in the Gulf of Maine area. Fishery Bulletin **97**:842-860.

Hutchings, J. A. 1999. Influence of growth and survival costs of reproduction on Atlantic cod, *Gadus morhua*, population groth rate. Canadian Journal of Fisheries and Aquatic Sciences **56**:1612-1623.

Hutchings, J. A. 2000a. Collapse and recovery of marine fishes. Nature **406**:882-885.

Hutchings, J. A. 2000b. Numerical assessment in the front seat, ecology and evolution in the back seat: time to change drivers in fisheries and aquatic sciences? Marine Ecology Progress Series **208**:299-313.

Hutchings, J. A., and R. A. Myers. 1993. Effect of age on the seasonality of maturation and spawning of Atlantic cod, *Gadus morhua*, in the Northwest Atlantic. Canadian Journal of Fisheries and Aquatic Sciences **50**:2468-2474.

Jørgensen, T. 1990. Long-term changes in age at sexual maturity of Northeast Arctic cod (*Gadus morhua* L.). Journal du Conseil international pour l'Exploitation de la Mer **46**:235-248.

Law, R. 2000. Fishing, selection, and phenotypic evolution. ICES Journal of Marine Science **57**:659-668.

Law, R., and D. R. Grey. 1989. Evolution of yields from populations with age-specific cropping. Evolutionary Ecology **3**:343-359.

Manly, F. J. 1991. Randomization, bootstrap and Monte Carlo methods in biology. Chapman & Hall, London.

Morgan, J., J. Burnett, and E. Aro. 1999. Variations in maturation, growth, condition and spawning stock biomass production in groundfish. Journal of Northwest Atlantic Fishery Science **25**:1-248.

Morgan, M. J., and E. B. Colbourne. 1999. Variation in maturity-at-age and size in three populations of american plaice. ICES Journal of Marine Science **56**:673-688.

Mousseau, T. A., and D. A. Roff. 1987. Natural selection and the heritability of fitness components. Heredity **59**:181-197.

Murawski, S. A., P. J. Rago, and E. A. Trippel. 2001. Impacts of demographic variation in spawning on reference points for fishery management. ICES Journal of Marine Science **58**:1002-1014.

Myers, R. A., N. J. Barrowman, J. M. Hoenig, and Z. Qu. 1996. The collapse of cod in Eastern Canada: the evidence from tagging data. ICES Journal of Marine Science **53**:629-640.

Myers, R. A., and N. G. Cadigan. 1995. Was an increase in natural mortality responsible for the collapse of northern cod? Canadian Journal of Fisheries and Aquatic Sciences **52**:1274-1285.

O'Brien, L. 1999. Factors influencing the rate of sexual maturity and the effect on spawning stock for Georges Bank and Gulf of Maine Atlantic cod *Gadus morhua* stocks. Journal of Northwest Atlantic Fishery Science **25**:179-203.

O'Brien, L., and N. J. Munroe. 2000. Assessment of the Georges Bank Atlantic cod stock for 2000. NEFSC, NEFSC Ref. Doc. 2000-17. 51 p.

Pigliucci, M., P. Diiorio, and C. D. Schlichting. 1997. Phenotypic plasticity of growth trajectories in two species of *Lobelia* in response to nutrient availability. Journal of Ecology **85**:265-276.

Purchase, C. F., and J. A. Brown. 2001. Stock-specific changes in growth rates, food conversion efficiencies, and energy allocation in response to temperature change in juvenile Atlantic cod. Journal of Fish Biology **58**:36-52.

Ratner, S., and R. Lande. 2001. Demographic and evolutionary responses to selective harvesting in poulations with discrete generations. Ecology **82**:3093-3104.

Reznick, D. A., H. Bryga, and J. A. Endler. 1990. Experimentally induced life-history evolution in a natural population. Nature **346**:357-359.

Reznick, D. N., F. H. Shaw, F. H. Rodd, and R. G. Shaw. 1997. Evaluation of the rate of evolution in natural populations of guppies (*Poecilia reticulata*). Science **275**:1934-1937.

Ricker, W. E. 1981. Changes in the average size and average age of pacific salmon. Canadian Journal of Fisheries and Aquatic Sciences **38**:1636-1656.

Rijnsdorp, A. D. 1989. Maturation of male and female North Sea plaice (*Pleuronectes platessa* L.). Journal du Conseil international pour l'Exploitation de la Mer **46**:35-51.

Rijnsdorp, A. D. 1993a. Fisheries as a large-scale experiment on life-history evolution: disentangling phenotypic and genetic effects in changes in maturation and reproduction of North Sea Plaice, *Pleuronectes platessa* L. Oecologia **96**:391-401.

Rijnsdorp, A. D. 1993b. Relationship between juvenile growth and the onset of sexual maturity of female North sea plaice, *Pleuronectes platessa*. Canadian Journal of Fisheries and Aquatic Sciences **50**:1617-1631.

Roff, D. A. 1992. The evolution of life histories. Theory and analysis. Chapman & Hall, New York.

Rohr, D. H. 1997. Demographic and life-history variation in two proximate populations of viviparous shrink separated by a steep altitudinal. Journal of Animal Eclogy **66**:567-578.

Ruzzante, D. E., C. T. Taggart, D. Cook, and S. Goddard. 1995. Genetic differentiation between inshore and offshore Atlantic cod (*Gadus morhua*) off Newfoundland: microsatellite DNA variation and antifreeze level. Canadian Journal of Fisheries and Aquatic Sciences **53**:634-645.

Scheiner, S. M. 1993. Genetics and evolution of plasticity. Annual Review of Ecology and Systematics **24**:35-68.

Serchuk, F. M., M. D. Grosslein, R. G. Lough, D. G. Mountain, and L. O'Brien. 1994. Fishery and environmental factors affecting trends and fluctuations in the Georges Bank and Gulf of Maine Atlantic cod stocks: an overview. ICES Marine Science Symposium **198**:77-109.

Sorci, G., J. Clobert, and S. Belichon. 1996. Phenotypic plasticity of growth and survival in the common lizard *Lacerta vivipara*. Journal of Animal Ecology **65**:781-790.

Stearns, S. 1989. The evolutionary significance of phenotypic plasticity. BioScience **39**:436-445.

Stearns, S. C. 1992. The evolution of life histories. Oxford University Press, Oxford.

Stearns, S. C., and J. C. Koella. 1986. The evolution of phenotypic plasticity in life-history traits: predictions of reaction norms for age and size at maturity. Evolution **40**:893-913.

Stokes, K., and R. Law. 2000. Fishing as an evolutionary force. Marine Ecology Progress Series **208**:299-313.

Sultan, S. E. 1996. Phenotypic plasticity for offspring traits in *Polygonom persicaria*. Ecology **77**:1791-1807.

Thompson, J. N. 1998. Rapid evolution as an ecological process. Trends in Ecology and Evolution **13**:329-332.

Trippel, E. A. 1998. Egg size and viability and seasonal offspring production of young Atlantic cod. Transactions of the American Fisheries Society **127**:339-359.

Trippel, E. A. 1999. Estimation of stock reproductive potential: history and challenges for Canadian Atlantic gadoid stock assessments. Journal of Northwest Atlantic Fishery Science **25**:61-81.

Trippel, E. A., M. J. Morgan, A. Fréchet, C. Rollet, A. Sinclair, C. Annand, D. Beanlands, and L. Brown. 1997. Changes in age and lenght at sexual maturity of Norwest Atlantic cod, haddock and pollock stocks, 1972-1995. Canadian Technical Report of Fisheries and Aquatic Sciences **2157**:1-120.

Wootton, R. J. 1998. Ecology of teleost fishes. Kluver Academic Publishers, Dordrecht.

Ylikarjula, J., M. Heino, and U. Dieckmann. 1999. Ecology and adaptation of stunted growth in fish. Evolutionary Ecology **13**:433-453.