CM 2002/Y:03

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

S. Barot[1 4], M. Heino[2 1], L. O'Brien[3] and U. Dieckmann[1]
[1] Corresponding author, Adaptive Dynamics Network, International Institute for Applied Systems Analysis. A-2361 Laxenburg, Austria. E-mail: sebastien.barot@bondy.ird.fr.
[2] Institute of Marine Research, Po Box 1870 Nordnes, N-5817 Bergen, Norway
[3] National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, Massachusetts 02543, USA
[4] Present address: Laboratoire d'Etude des Sols Tropicaux, IRD, 32 Avenue Henri Varagnat, 93143 Bondy Cedex, France

Abstract: Average age and size at maturation have decreased in many commercially exploited fish stocks during the last few 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, and to disentangle evolutionary changes and phenotypic plasticity, we use a reaction norm based approach. A method is presented to estimate the reaction norm for age and size at maturation from a type of data commonly gathered for the management of fisheries. This method is applied to data on Georges Bank and Gulf of Maine Atlantic cod stocks. Maturation reaction norms in these stocks have shifted significantly downwards, i.e. there has been a tendency to mature earlier at smaller size. These findings support the hypothesis of an evolutionary trend, probably caused by the selective removal of larger fish and high fishing mortality rates. Consequences of such an evolutionary process for the sustainability of the fishery are discussed.

Keywords: cod, fisheries, logistic regression, maturation reaction norm, maturity ogive, phenotypic plasticity, life-history evolution

INTRODUCTION

Life-history parameters, such as age and size at first reproduction, survival rate or the number of offspring, vary in space and time for a given species (Roff 1992, Stearns 1992). They are partially genetically determined and evolve according to selective pressures. They also depend on environmental variations expressed through phenotypic plasticity, which often has an adaptive value of 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 these direct and indirect environmental effects, i.e. the respective influence of phenotypic plasticity and evolution. Short-term experiments have assessed 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 the long-term trend in a life-history trait and attempted to infer the causes of this trend, mostly because suitable data is seldom available.

We contribute to filling this gap by taking advantage of time-series collected to support the management of commercially exploited fish stocks (Hilborn and Walters 1992). We chose to study age and size at maturation because fitness is very sensitive to these life-history parameters (Roff 1992, Stearns 1992). For individuals, they influence the potential number of reproductive events. At the population level, they determine the size and age distributions of reproducing individuals, and influence the 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 strongly influence population dynamics and potential yields of commercially fished stocks. Moreover, 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 purely phenotypic 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 a few decades? This question is interesting per se to improve our understanding of life-history evolution. This issue is also important to assess the sustainability of fisheries and current management strategies.

In fisheries science, maturation dynamics is usually described by two indices: the age at which 50% of individuals are mature and the length at which 50% of individuals are mature, A_{50} and L_{50} (Jørgensen 1990, Chen and Paloheimo 1994, Morgan and Colbourne 1999, O'Brien 1999). In many fish stocks these indices suggest that maturation processes 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). This trend could be due to purely phenotypic changes resulting from any long-term trend in the environment, e.g. the temperature or the population density, the latter influencing many relevant parameters such as food availability. Population density is actually a good candidate because it has decreased in many stocks due to high fishing mortality rates. Reduced population density is likely to partially relax density-dependent effects that slow growth and delay maturation. The alternative explanation for the trend in A_{50} and L_{50} is that high fishing mortality rates could also cause genetic decreases in age and size at maturation, because fishing always alters the pattern of size- and age-dependent mortality and can select for specific combinations of age and size at maturation (Rijnsdorp 1993a, Law 2000, Stokes and Law 2000, Ratner and Lande 2001). These two explanations are not mutually exclusive.

The indices A_{50} and L_{50} are estimated using a logistic regression to predict the probability of being mature as a function of age or size. Because the curves describing this probability are called maturity ogives, we will refer to the approach using these indices to characterize changes in maturation as the ogive approach. A_{50} and L_{50} indices describe the maturation process only partially. First, theoretical models show that maturation depends on both age and size (Roff 1992, Stearns 1992) and 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 analysis. Second, these indices describe the probability of being mature, which depends not only on maturation processes, but also on survival and growth variations 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 maturation processes have not changed (Heino, et al. 2002b). Characterizing maturation processes with the probability of maturing addresses these problems. When calculated as a function of both age and size the probability is by definition conditioned on age and size, and therefore allows the characterization of maturation processes independently from the processes of growth and survival. This probability corresponds to the probabilistic extension of

the classical reaction norm for age and size at maturation (Stearns 1992, Heino, et al. 2002b). 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. Figures often only display the size at which the probability of maturing is 50%, the so-called reaction norm midpoints, which are usually plotted as a function of age. 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 rate 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, if two populations have different maturation reaction norms, we can assume that they are genetically different. 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.

Barot et al. (2002) have shown how to estimate the reaction norm for age and size at maturation when data on both mature and immature individuals are collected annually. The method enables the estimation of the probability of maturing at age and size for individual cohorts. Unfortunately, it requires a large sample size for the estimations to be robust. Our goal is twofold. First, we describe an improvement in the estimation method of Barot et al. (2002). This improvement allows for the use of smaller sample sizes by assuming a common pattern in the shape of the reaction norm of the different cohorts. Second, we estimate the reaction norms for age and size at maturation for two Atlantic cod (*Gadus morhua*) stocks. We use artificial data to test the robustness of the method and to facilitate the interpretation of the results. We then discuss the biological interpretation and the fishery implications of these results.

MATERIAL AND METHODS

Biological and environmental data

The method used 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. 2002). We examine data for two stocks of Atlantic cod in the Northwest Atlantic, Georges Bank and the Gulf of Maine (hereafter GB and GM), of the Northwest Atlantic. The exploitation rate on both stocks has increased over the last four decades, due to the increased effort of the distant-water fleets during 1960-1970, and the subsequent increased effort of both due USA and Canadian fisheries. The data is obtained from a spring bottom trawl survey conducted since 1968 by a research vessel of the Northeast Fisheries Science Center (Azarowitz 1981, O'Brien 1999). Sampling is random but stratified by length.

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 research vessel bottom trawl survey (O'Brien 1999). The second variable, the spring stratified mean weight per tow (kg), is an index of the cod stock biomass. It allows assessing density-dependent effects on

maturation, that could, for example, be due to a decline in food availability when the stock biomass is high (O'Brien 1999, O'Brien and Munroe 2000).

General description of the estimation method

The probability of maturing at age *a* and size *s, m*(*a*,*s*), can be calculated from estimations of the probabilities of being mature at age and size, *o*(*a*,*s*), and from estimations of the mean annual growth at age, Δs (Barot, et al. 2002):

$$m(a,s) = \frac{o(a,s) - o(a-1, s-\Delta s)}{1 - o(a-1, s-\Delta s)} . \quad (1)$$

This equation is strictly valid only under the assumption that immature and mature individuals have, within an age and size class, the same survival and growth rates. However, we have shown that the estimation is robust to violations of these assumptions (Barot, et al. 2002). The full estimation method involves four steps: (1) *o*(*a*,*s*) is estimated through a logistic regression. (2) Δs is estimated as the difference between the mean size at age for two consecutive ages. (3) *m*(*a*,*s*) is computed using Eq. 1. (4) A facultative step is to summarize the array of *m*(*a*,*s*) values 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 used to compare easily 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 it 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 as a last step, because the logistic curve may not always fit well the estimated probability of maturing.

The estimation procedure outlined above has been applied independently to individual cohorts (Barot, et al. 2002). To estimate robustly reaction norm midpoints, at minimum 100 individuals must be sampled for each considered age and cohort (Barot, et al. 2002). 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 simultaneously information on all available cohorts by describing the maturity ogive with a single logistic model.

Estimation of annual growth

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

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) using the logit link function:

$$\text{logit}(o(a,s,c)) = \alpha_0 + \alpha_c + \alpha_a + \alpha_{a,c} + \beta_0 s + \beta_c s + \beta_a s + \beta_{a,c} s . \quad (2)$$

Parameter α_0 is the general intercept term and β_0 is the average effect of size. The constants α_c, α_a and $\alpha_{a,c}$ express the effects of age and cohort considered as factors, as well as their interaction, while the constants β_c, β_a and $\beta_{a,c}$ express the age- and cohort-dependent effects of size considered as a variate. This statistical model is a full model: all possible interactions between the

three independent variables, age, size, and cohort, are taken into account. This full model corresponds to the independent estimation of maturity ogives for each cohort and each age.

This approach should lead to the least-biased results, but it is not robust when sample sizes are low (Barot, et al. 2002): when too many parameters are estimated relative to the sample size, standards errors of the estimated parameters increase, and parameters 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. The final model used was determined after preliminary analyses were conducted, and is presented in the results section.

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. When, in a given cohort, n_a individuals have been sampled at age a, n_a individuals were 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. 2002). The process is repeated a 1000 times and the resulting distribution of the estimated midpoints 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 maturation of females and males would only arise by chance. Such a data set is used to estimate independently the maturity ogives for males and females. The probability of maturing is then computed 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:

$$\text{logit}\left(m(s)\right) = \alpha_0 + \alpha_c + \alpha_{sex} + \beta_0 s \,.$$

This randomization procedure is repeated 1000 times and the likelihood-ratio χ^2 statistic computed and collected to test for the sex effect. The same calculations are applied to the observed 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 observed 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:

$$\text{logit}\left(m(s)\right) = \alpha_0 + \alpha_c + \alpha_{stock} + \beta_0 s \quad \text{and}$$

$$\text{logit}\left(m(s)\right) = \alpha_0 + \beta_0 s + \beta_1 c \,.$$

In this last 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:

$$\text{logit}\left(m(s)\right) = \alpha_0 + \alpha_a + \beta_0 s + \beta_a c + \gamma_0 \; c\, s \,.$$

RESULTS

Environmental variations

There is a significant long-term trend in the biomass of both the GB and the GM cod stocks (Fig. 1, 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 fluctuates but exhibits no long-term trend (linear regression, GB, $P>0.05$; GM, $P>0.05$).

Growth

Growth varies from year to year but no long-term temporal trend is evident for ages 1-5 for female and male cod on GB and in the GM (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 GB and 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 o the results' below) showed that for the GB and GM data sets, the sample size at age is too low to analyse the interaction between age and size using both age and size as factors (discrete variables). Consequently the full statistical model (Eq. 2) 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:

$$\text{logit}\left(o\left(a,s,c\right)\right) = \alpha_0 + \alpha_c + \beta_c a + \beta_0 s \,.$$

The consequences of this statistical 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 interaction between size and age or size and cohort was 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. 2002).

Maturation reaction norms

Midpoints of the maturation reaction norms and the corresponding confidence intervals were estimated for each cohort, separately for male and females of both GB and GM cod stocks. As an example, Fig. 3 displays the female reaction norms for age and size at maturation assessed for the 1980 cohort of 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 growth curve. This results simply from the fact that more maturation data is available for these ages.

Reaction norms, averaged over five years periods and estimated by sex and by stock indicate a temporal trend towards maturation at smaller size (vertical shift of the reaction norms, Fig. 4). This pattern is stronger for GB than for GM cod. The existence of this trend is confirmed by the results of randomization (Table 1). This result is also obvious looking at the values of the midpoints for all cohorts without any pooling (Fig. 5), despite the large short-term variations in the midpoints.

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, randomization tests for an age effect, $P<0.05$, in these two cases predicted midpoints decrease from age 1 to 5), i.e. at the same size old individuals have higher probabilities of maturing than young ones. There is a trend, clearer for GB, towards an increase of the reaction norm slope (Figs. 4 and 5). This means that for the earlier cohorts age tends not to influence the probability of maturing, while for the more recent ones the probability of maturing at a given size increases with age.

It must be emphasized that the maturation reaction norm is in fact constituted by the set of curves describing the probability of maturing as a function of age and size. Fig. 4 and 5 only display the reaction norm midpoints, i.e. the sizes at which the probability of maturing is 50 %. This is very useful to compare the probability of maturing between ages and populations, but this does not mean that fish really mature at these sizes. For example, it is obvious, looking at the mean size at age on Fig. 3, that five-year-old cods are on the average taller than the reaction norm midpoint (45 cm). Thus the reaction norm means that immature five-year-old cods have a probability of maturing much larger than 75 % and mature on the average when they are about 70 cm long. Conversely, the reaction norm indicates that the probability of maturing at age one is much smaller than 25 % (in fact, this probability is nearly zero) because at this age cods are on the average 25 cm long.

Significant differences between sexes and stocks are revealed by randomization tests (Table 1). First, they show that 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). Because the reaction norms are nearly horizontal or are slightly tilted downwards for older ages that means that females tend to mature at smaller sizes, and slightly younger ages than males. Second, the probabilities of maturing at age and size tend for both sexes to be higher for GB than for GM at age 2 (Table 1), while it is the reverse for age 4 and 5 for females (test not significant for males).

ROBUSTNESS OF THE RESULTS

The estimation method to calculate the probability of maturing has been shown to be robust to the hypothesis that growth and survival rates are similar for juvenile and mature individuals at a given size (Barot, et al. 2002). 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 using smaller sample sizes. However, model simplifications might result in biases. 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 environmental conditions and errors in determining the maturity status may have on the estimations. We do so by creating artificial data sets encompassing information on more than one cohort, using a priori theoretical probabilistic reaction norms (Barot, et al. 2002, Heino, et al. 2002a). These data sets can then be 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. (2002). Data is generated by 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 is unrealistic in one important way: it ignores the short-term environmental variability that might cause annual, cross-cohort anomalies in the tendency to mature. 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 a likely problem of data quality: some mature individuals are probably misclassified as mature, and vice versa. It might be difficult to distinguish an immature individual from a mature one for which the gonad is in a resting stage (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 should lead to an increase in the ogive steepness (the probability of being mature increases quicker with size), but the size at which 50% of individuals are mature is not biased. 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 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 correctly estimate 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 an other 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 have approximately the same size as the available data set for GB and GM cod stocks: 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 from the earlier to the later cohorts until it reaches a slope of -3 cm yr^{-1} (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 vertical temporal shift of the reaction norm 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. Estimations are not more biased for ages at which few individuals mature (ages 1 and 5 yrs): the error term in the estimated midpoints is 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. 6) for all but the higher

intensity of year effect (Y_{sd}=10) for which the bias is positive. For each age, errors in slope generally increase with the year effect (Y_{sd}).

When the successive cohorts are more and more tilted clockwise the temporal trend of the reaction norm midpoints is correctly detected (Table 3), at least when the year effect is not too high (Y_{sd}<6). When the year effect increases (up to Y_{sd}=8) the temporal trend in the reaction midpoints at age is qualitatively well estimated but not quantitatively (Fig. 6). For Y_{sd}<6 the estimated reaction norms are tilted as the actual ones, but it is no longer the case for higher year effects. Errors in the reaction norm slope (relation between midpoints and age) are low at least for Y_{sd}<8. Both the estimations of 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 data set quality. 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 valid, even if the year effect on maturation is strong.

DISCUSSION

Our first conclusion is methodological: it is possible to assess, even with small sample sizes, the long-term trend in maturation reaction norms if data on both immature and mature individuals is available. Consequently, the methodology is in place to use 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, for GM and GB cod stocks, a shift of the maturation reaction norm towards lower ages and sizes at maturation. 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 detected with the maturity ogive approach.

Maturity ogive vs. maturation reaction norms

Variations in maturation of the GB and GM cod stocks have earlier been studied using estimations of the probability of being mature, i.e. the maturity ogives (O'Brien 1999). We examine the same problem using the reaction norm approach. Results of both approaches are partially consistent. First, they both reveal a long-term trend towards lower ages and sizes of reproducing individuals. Second, they show that individual cods tend to start reproducing at smaller sizes at age on GB than in the GM at age two and the reverse for older ages. In apparent contradiction with former results using the maturity ogive approach for other cod stocks (Beacham 1983, Trippel, et al. 1997, Ajiad, et al. 1999), males tend to start reproducing at higher ages and larger sizes than females. It must be emphasized that the reaction norm and the maturity ogive approaches bring forwards qualitatively different information. The probability of being mature depends not only on maturation processes, but also on variations in the growth and survival rates, while the probability of maturing only depends on maturation. Consequently, it is easier to interpret variations in the probability of maturing than in the probability of being mature.

Temporal or spatial variations in the probability of maturing are potentially due to both genetic and to environmental variations. However, we have used the reaction norm approach for age and size at maturation that focuses on maturation probability conditioned on age and size, so that these two nonexclusive possibilities are partially disentangled. Taken together, by definition of this type of reaction norm (Stearns and Koella 1986, Heino, et al. 2002b, Heino, et al. 2002a), differences between reaction norms are due to genetic differences, or to environmental variations

that do not result in growth rate variations, because different growth trajectories would pass through the reaction norm in different points.

We acknowledge that the denomination of reaction norm for maturation reaction norms can be disputed because environmental variations are not explicit, because maturations can be determined by other variables that are not taken into account (weight, liver index) and because growth variations are not only environmentally determined. However, size at age has been proved to be very sensitive to environmental variations both using tank experiments and field data (Wootton 1998, Imsland and Jónsdóttir 2002) and long-term time series are seldom available for other relevant variables. Moreover, it is likely that the growth of an individual is partly determined by its genes (Conover and Schultz 1995, Wootton 1998, Imsland and Jónsdóttir 2002), but that should not impede the conclusion that two different maturation reaction norms support the idea of two different genotypes. This conclusion should only be hindered if the evolution of growth is strongly linked to the evolution of maturation through some common genes that act both on sexual maturation and growth, which is unlikely.

Interpretation of maturation reaction norms

How to interpret the maturation difference between males and females? The difference between male and female maturation reaction norms is unambiguous to interpret because males and females of a given stock experience the same environment. The differences in probability of maturing must therefore 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 female reproduction requires usually more energy than male reproduction, so that males have evolved smaller sizes and ages at maturation (Stearns 1992). Nevertheless, the observed difference between sexes is very small (Fig. 4 and Fig. 5) and only significant in two cases (age 2 and 3, Table 1). The reaction norm approach could be applied to other cod stocks to check if the differences between sexes found with the maturity ogive approach (Beacham 1983, Trippel, et al. 1997, Ajiad, et al. 1999) correspond actually 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 maturation difference between males and females. This is due to the fact that environmental conditions vary in time (Fig. 1), and that the two stocks experience different environments (O'Brien 1999). This alone could explain, through phenotypic plasticity, the temporal trend in the reaction norm and maturation differences between the two stocks. However, there is no long-term trend in growth rate, while there is a significant difference between the growth rates of the two stocks.

How to interpret the temporal trend in the maturation reaction norms? Two non-exclusive hypotheses can explain this trend. A first hypothesis is that selective pressure by the fishery has caused evolution towards low age and size at maturation. A second hypothesis is that the temporal trend is due to phenotypic plasticity. The maturation trend should then be explained by a long-term trend in the environment; these two trends would necessarily be parallel. In this context, stock biomass is a good environmental candidate variable because it displays a decreasing trend since 1970 (Fig. 1). Moreover, stock biomass is a relevant variable because it may influence cod life-history through density-dependant processes. For example, food availability may increase when biomass is low. However, this phenotypic plasticity scenario is improbable. Tank experiments and field data show that growth is sensitive to most environmental variables, so that growth is likely to be sensitive to any environmental variable that could influence maturation through phenotypic plasticity (Wootton 1998a). Consequently, if the temporal trend in maturation was due to phenotypic plasticity, we should also observe a temporal trend in the growth rate, which is not the case.

Of the other environmental variables that could explain the observed maturation trend through phenotypic plasticity, water temperature can be ruled out because this variable does not show any long-term trend (Fig. 1). Because no trend in growth has been detected, other candidate variables would need to influence maturation without influencing size-at-age at the same time. Though such a variable could exist, we are not aware of any evidence pointing to this direction. For all these reasons, the long-term trend in maturation reaction norms and the corresponding decline in age and size at maturation are probably mostly due to a genetic change.

A decrease in age and size at maturity has been detected in many fish stocks using the maturity ogive approach: American plaice (Morgan and Colbourne 1999), North sea plaice (Rijnsdorp 1989, Rijnsdorp 1993a) and cod (Jørgensen 1990, Trippel, et al. 1997). The next step to understand these declines would be to check, using maturation reaction norms, whether they are really due to changes in the maturation process. It would be interesting to compare the reaction norm trends of stocks that have different fishing histories. It would be particularly useful to apply the reaction norm approach to a stock for which the total biomass has first decreased, but is now increasing due to some fishing restrictions. If age and size at maturation have constantly been decreasing, very few doubts would remain on the fact that the maturation has truly evolved due to the fishing pressure.

How to interpret the maturation differences between the two stocks? The ogive approach suggested clearly that cod matures earlier and at smaller sizes on GB than in the GM (O'Brien 1999). The reaction norm approach gives more balanced results: of four significant tests out of ten, two indicate that maturation is at a smaller size on GB than in the GM, and two indicate the reverse. The discrepancy between the two approaches is due to the fact that only the reaction norm approach corrects for differences in growth rates, which are actually higher on GB than in the GM (Fig. 3). This means that the difference in maturation between the two stocks is actually less important than predicted earlier, and that this difference might be reversed for older ages. Nevertheless, GM and GB cod stocks experience different environmental conditions (O'Brien 1999). GB constitutes a highly productive shoal averaging 50 m in depth, while the GM is a deeper area with an average depth of 150 m. Moreover, the autumn water temperature is higher for GB than for the GM. These differences correspond to relevant traits of the physical environment of a cod because they result in faster growth rates on GB than in the GM. It is thus possible that these environmental differences are important enough to lead to purely phenotypic changes in age and size at maturation, which are not taken into account by the reaction norm approach. The maturation differences between the two stocks are due, at least partially, to phenotypic plasticity. Yet "common garden" experiments (fish of different stocks are kept in the same environmental conditions) 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 shown that geographically adjacent cod stocks can be genetically different (Ruzzante, et al. 1995). It cannot be excluded that maturation differences between GB and GM stocks have also a genetic component.

Can fishing pressure lead to rapid evolution?

Although we have not strictly proved that age and size at maturation have genetically changed, our analyses support this interpretation. It has been widely recognized since Borisov's (1978) and Ricker's (Ricker 1981) pioneering work that such an evolution of a life-history trait could be due to the selective pressure exerted by fishing (Ylikarjula, et al. 1999, Hutchings 2000b, Law 2000, Stokes and Law 2000, Ratner and Lande 2001). However, this idea has been mostly developed using verbal arguments and theoretical models. It is one of the first times that the evolution of a life-history parameter at the scale of several decades is suggested so clearly with field data. Fishing can be considered here as a long-term experiment, while former studies have used a real experimental approach (Reznick, et al. 1990, Reznick, et al. 1997).

Because fishing gears are size-selective, and fishing is age-selective, at least through this size dependence, it is generally accepted that fishing represents a selective pressure for life-history traits. 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? On the one hand, 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). On the other hand, 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 is commonly so high that it results in very high selection differentials (Law 2000), and fishing mortality has been indeed very high for Gulf of Maine and Georges Bank cod stocks (O'Brien 1999, O'Brien and Munroe 2000, Mayo, et al. 2002).

High evolutionary rate require not only 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 (around 0.31, mean of height fish breeding experiments) have been found for age at maturation (Law 2000), and such values should not preclude quick response to selection. It must be noted that assessing the heritability of maturation parameters in the wild is the prerequisite for estimating how quickly such traits can evolve 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

We have shown that our analyses are relevant to tackle strictly evolutionary issues. However, our results and forthcoming studies on temporal trend in the probability of maturing are primordial to understanding the long-term dynamics of commercially fished species. Many fish stocks have collapsed, probably due to the combination 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 and their variability. These collapses have also resulted in more precautionary fishing strategies 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). This trend in age and size at maturation is likely to reduce indirectly the reproductive potential of the stock, leading to low recruitment rates. This would 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 potential and low age and size at maturity, constitutes one more 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 increase notably within a short period of time because selective pressures for small/large sizes, and young/old age at maturation are not symmetric (Law and Grey 1989, Heino 1998, Law 2000). Even such an extreme measure as closing a fishery is likely, at best, to create a moderate selection pressure for increased size and age at maturation. 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 would have a low chance of rebuilding. Our results support this pessimistic option. This suggests that designing sustainable fishing strategies requires not only taking into account all the factors influencing stocks dynamics itself, but also accounting for the evolutionary consequences of fishing pressure that constitute long-term feedback loops. Hence, we advise the use of fishing strategies that take into account the evolutionary effect of fishing, i.e. Darwinian fishing strategies (Law and Grey 1989, Heino 1998, Law 2000, Stokes and Law 2000).

ACKNOWLEDGEMENTS

This research has been supported by the European Research Training Network ModLife (Modern Life-History Theory and its Application to the Management of Natural Resources), funded through the Human Potential Programme of the European Commission (Contract HPRN-CT-2000-00051). M. Heino's work has been also funded by the Academy of Finland (grant 45928).

LITERATURE CITED

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

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

Barot, S., M. Heino, L. O'Brien, and U. Dieckmann. 2002. Estimation of reaction norm for age and size at maturity with missing first-time spawner data. submitted to Oecologia

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., and E. T. Schultz. 1995. Phenotypic similarity and the evolutionary significance of countergradient variation. Trends in Ecology and Evolution **10**:248-252.

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 in press

Heino, M., U. Dieckmann, and O. R. Godø. 2002b. Measuring probabilistic reaction norms for age and size at maturity. Evolution in press

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

Holzwarth, T., and D. Mountain. “Surface and bottom temperature distributions from the Northeast Fisheries Center spring and fall bottom trawl survey program, 1963-1987.” . Woods Hole: NOAA/National Marine Fisheries Service, 1990.

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.

Imsland, A. K., and Ó. D. B. Jónsdóttir. 2002. Is there a genetic basis to growth in Atlantic cod? Fish and Fisheries **3**:36-52.

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 Sciences **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.

Mayo, R. K., E. M. Thunberg, S. E. Wigley, and S. X. Cadrin. "The 2001 assessment of the Gulf of Maine Atlantic Cod stock." NEFSC, 2002.

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**:

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 George Bank and Gulf of Maine Atlantic cod *Gadus morhua* stocks. Journal of Northwest Atlantic Fisheries Science **25**:179-203.

O'Brien, L., and N. J. Munroe. "Assessment of the Georges Bank Atlantic cod stock for 2000." NEFSC, 2000.

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 on Ecology and Systematic **24**:35-68.

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 evolutionnary force. Marine Ecology Progress Series **208**:299-313.

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

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 Fisheries 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. 1998b. Growth. Pages 107-140 *in* Ecology of teleost fishes. Kluwer Academic Publishers. Dordrecht.

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

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 ♂<♀ 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	Concerned stock and 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	+**	+**	+**	+*	ns
	GM Females	ns	ns	+*	+**	+*

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 mean temporal trend, the mean absolute error in this trend and the mean absolute error in the reaction norm slope are displayed (see Table 2 caption). 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 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

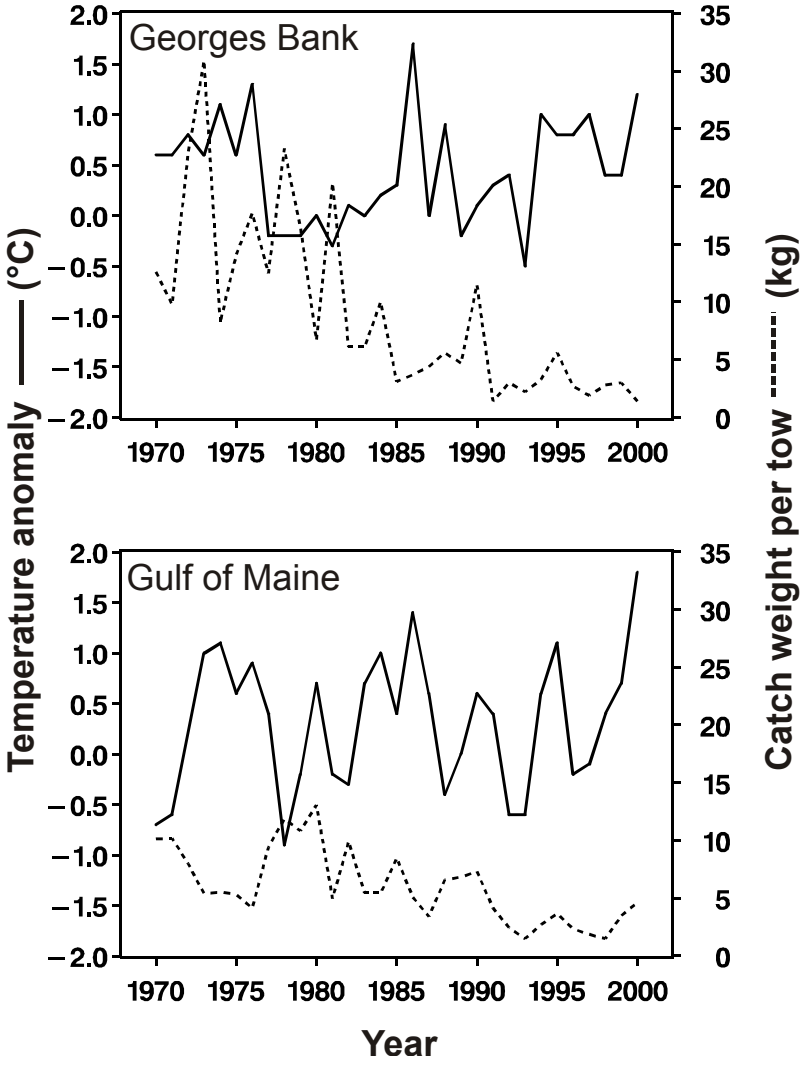

Figure 1. Temporal variations in two environmental indices: the mean weight per tow and the bottom spring temperature anomaly.

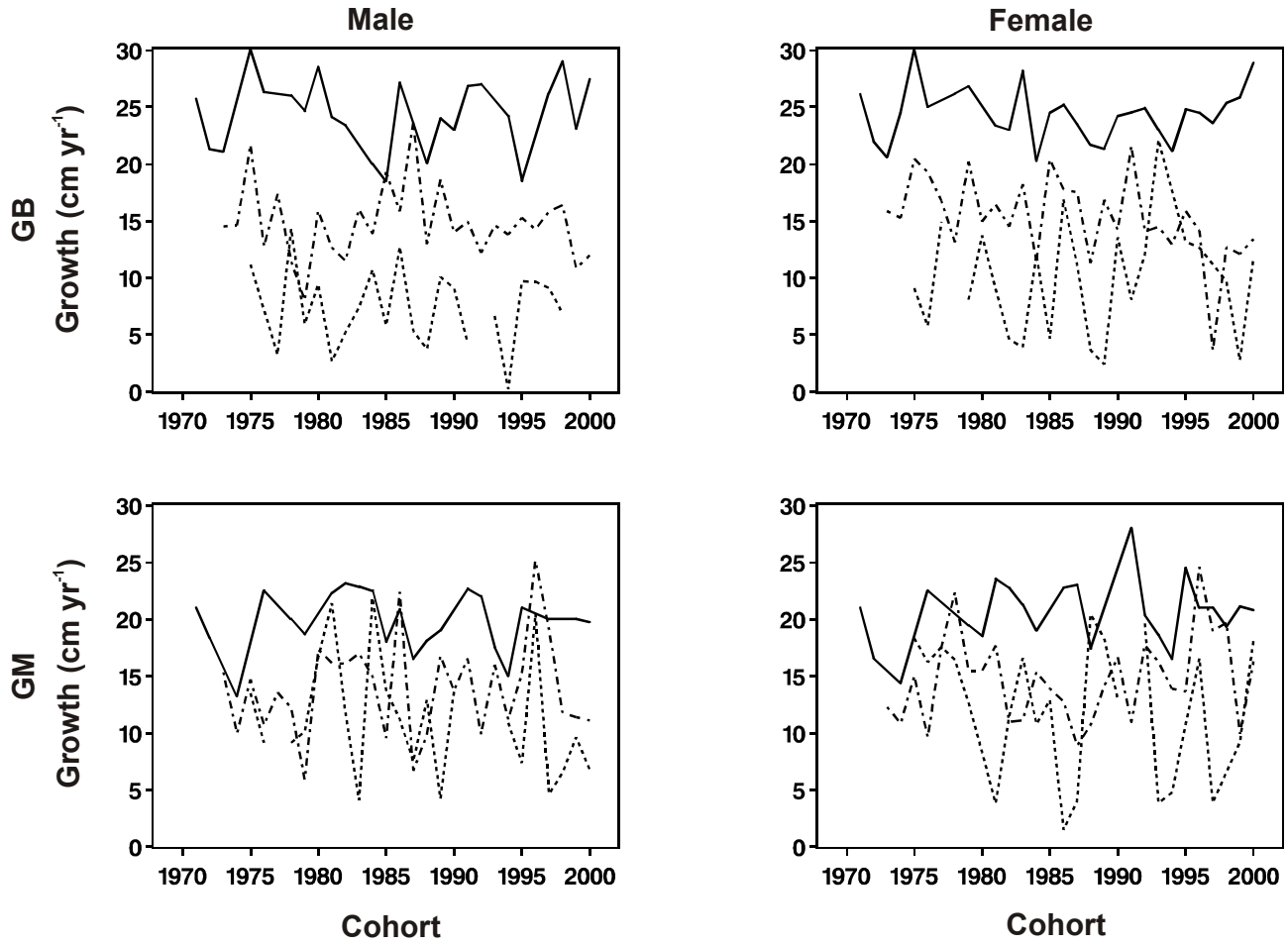

Figure 2. Annual growth increments estimated for each age and cohort. ——, age 1; — · —, age 3; ····, age 5. Missing points correspond to very low growth rates that were estimated to be negative. For clarity, curves for ages 2 and 4 are not displayed, but they present similar oscillations as ages 1, 3,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

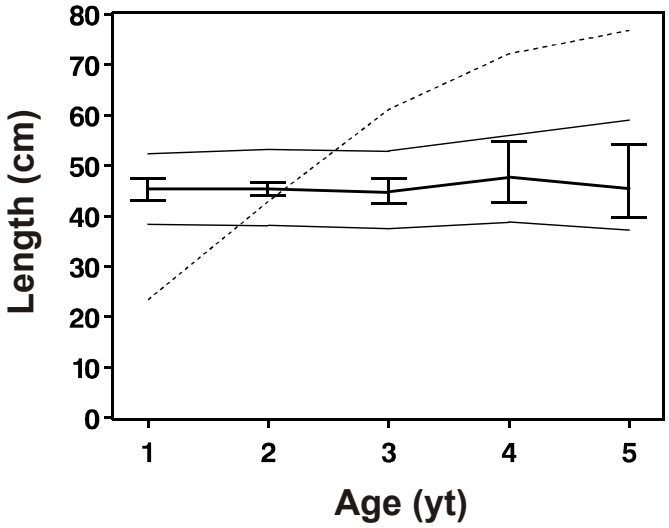

Figure 3. Single reaction norms (thick plain line) estimated for the females of the cohort 1980 and displayed with the inter-quartile range (thin continuous line) and bootstrapped confidence intervals for the midpoints (see method section). The mean size at age is also displayed (doted line).

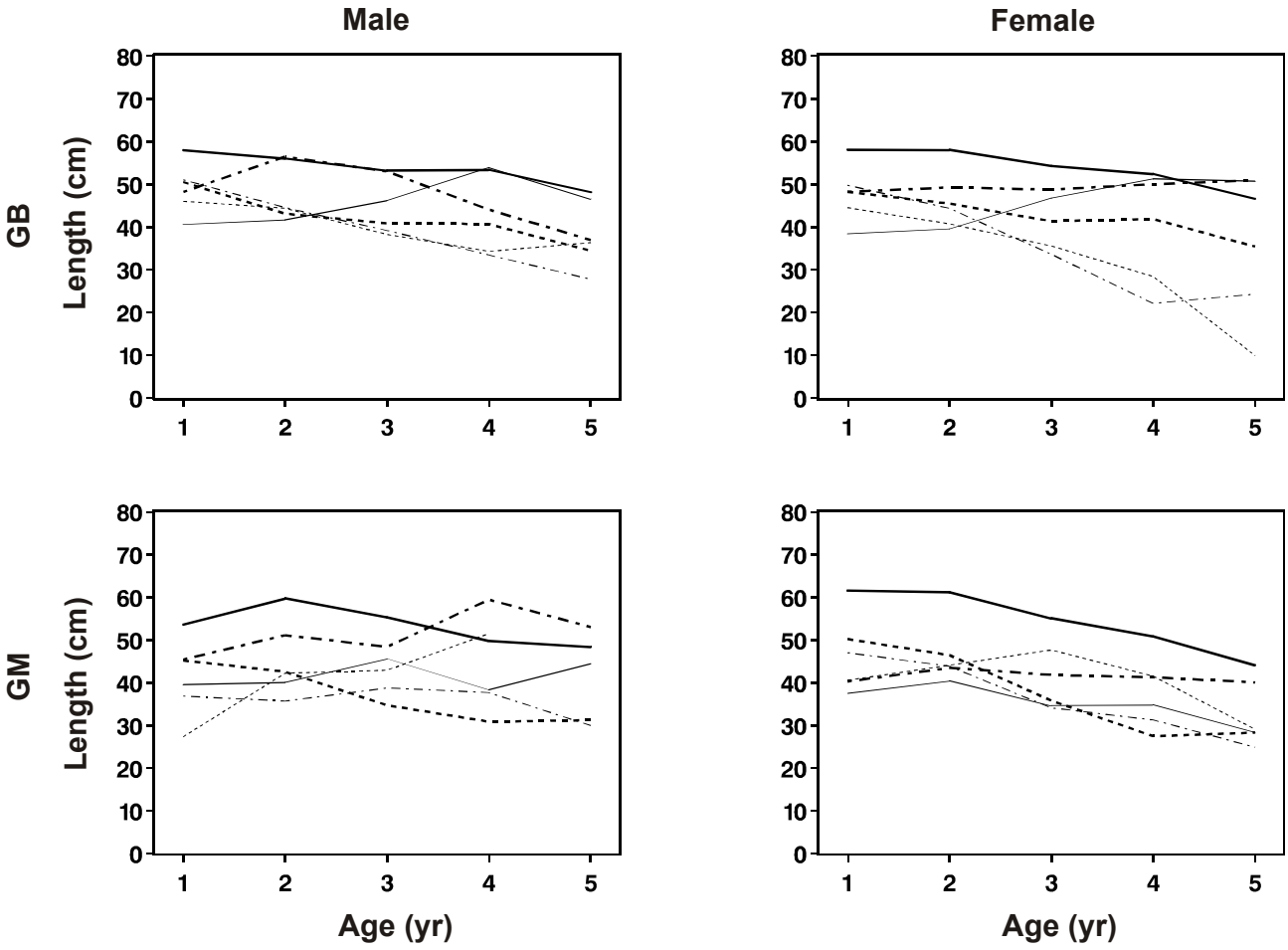

Figure 4. Reaction norms for age and size at maturation averaged over 5 years periods. Reaction norms have been estimated separately for males/females and the two stocks. ——, cohorts 1970-74; – · –, cohorts 1975-79; ····, cohorts 1980-84; ——, cohorts 1985-89; – · –, cohorts 1990-94; ·····, cohorts 1995-1999.

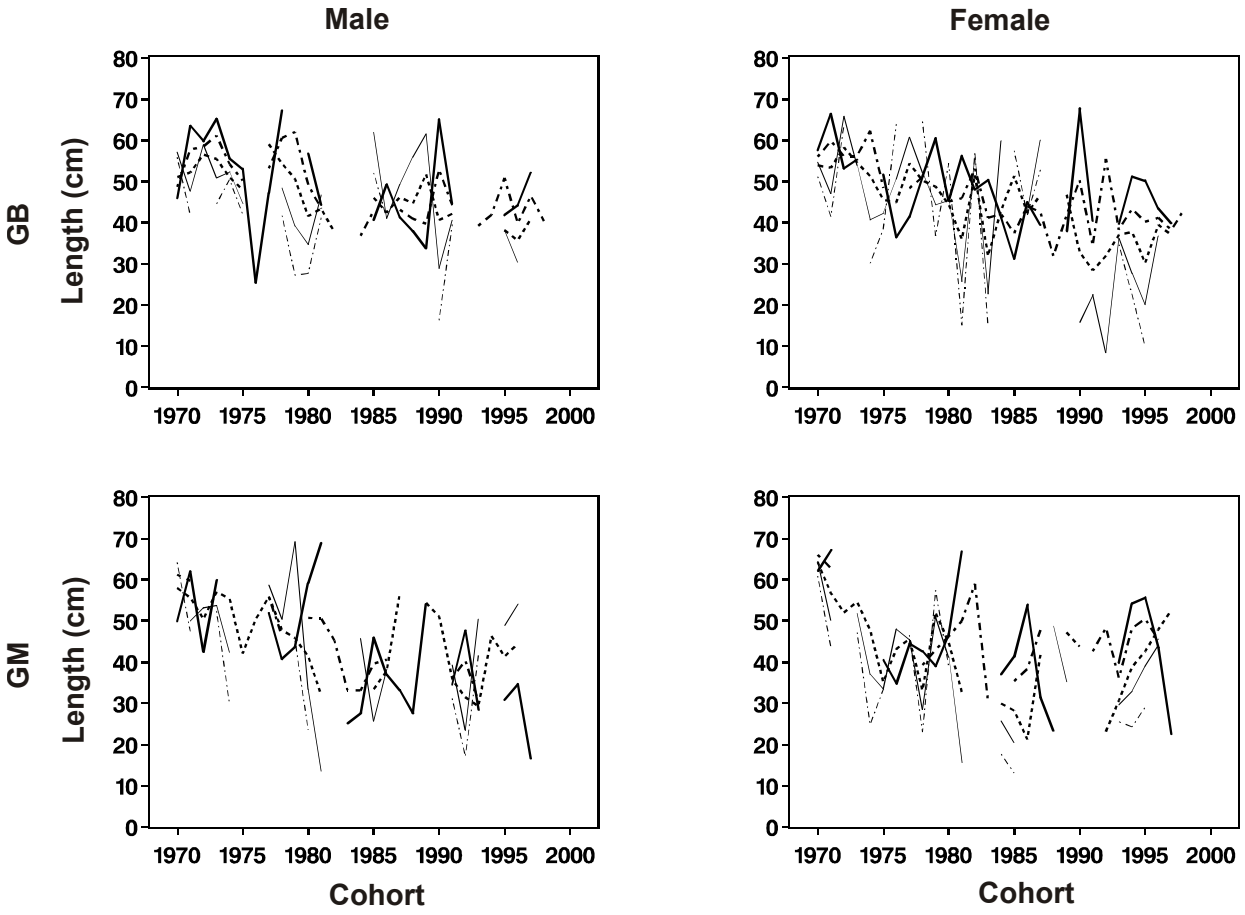

Figure 5. Temporal trend in the reaction norm for age and size at maturation midpoints. Each curve corresponds to a different age. Reaction norms have been estimated separately for males/females and the two stocks. Curves are not continuous because it was not possible to estimate some of the midpoints due to the sample size. ——— , age 1; — · — , age 2; ····· , age 3; ——— , age 4; — · — , age 5.

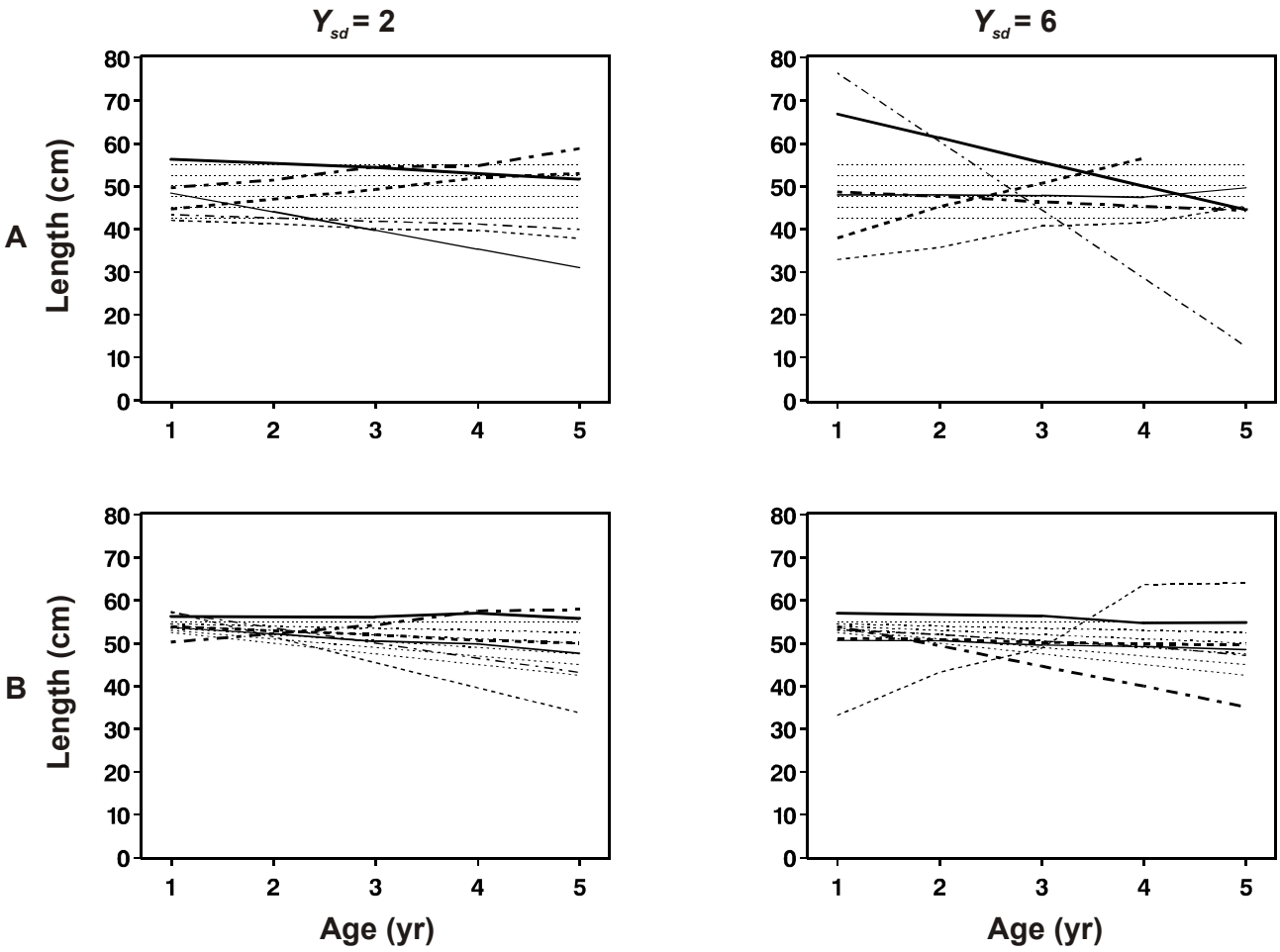

Figure 6. Graphic example of robustness tests. (A) For the first test (no age effect), theoretical reaction norms are horizontal and shift vertically. (B) For the second test (age effect), reaction norms of the successive cohort are more and more tilted clockwise. Five cases, corresponding to an increase in the year effect (Y_{sd}) on the reaction norm midpoints have been considered, and two cases are here displayed. Because the reaction norms of 30 cohorts can't be displayed on the same figure only the reaction norms of 6 evenly spaced cohorts are displayed. Thin doted lines, theoretical reaction norms; ——, cohort 0; — · —, cohort 5; ····, cohort 10; ——, cohort 15; — · —, cohort 20; ·····, cohort 25.