Fisheries-induced evolutionary changes in maturation reaction norms in North Sea sole *Solea solea*

Fabian M. Mollet*, Sarah B. M. Kraak, Adriaan D. Rijnsdorp

Wageningen Institute for Marine Resources and Ecological Studies (IMARES), PO Box 68, 1970 AB Ijmuiden, The Netherlands

ABSTRACT: Age and size at maturation decreased in several commercially exploited fish stocks, which, according to life history theory, may be due to fisheries-induced evolutionary change. However, the observed changes may also represent a plastic response to environmental variability. To disentangle phenotypic plasticity from evolutionary change, the probabilistic reaction norm approach was applied to 43 cohorts (1960 to 2002) of female sole *Solea solea* from market samples. The reaction norm for age and size at first maturation has significantly shifted towards younger age and smaller size. Size at 50% probability of maturation at Age 3 decreased from 28.6 cm (251 g) to 24.6 cm (128 g). This change was even stronger when condition was included as a third dimension in the reaction norm estimation. The influence of alternative factors was tested on the population level by regression of reaction norm midpoints on annual estimates of condition, temperature and competitive biomass. Although effects of temperature and competitive biomass were significant, the variation in the midpoints was best explained by the decreasing time trend. Therefore, the results provide strong evidence for a fisheries-induced evolutionary change in the onset of sexual maturity.

KEY WORDS: Life history evolution · Probabilistic maturation reaction norm · Phenotypic plasticity · Growth · Condition

INTRODUCTION

Changes in life history traits such as age and size at first maturation have been reported in several commercially exploited fish stocks (Jørgensen 1990, Rijnsdorp 1993a). These changes may be due to either (1) phenotypic plasticity in response to environmental change or (2) fisheries-induced evolutionary change (Stokes et al. 1993). If mortality increases, earlier maturation and higher reproductive investment will result in a higher fitness (Law 2000). These predictions have been confirmed experimentally (D. A. Reznick et al. 1990, Conover & Munch 2002); the challenge is to examine how these processes occur in exploited wild populations.

Evolutionary changes, however, may be overshadowed by phenotypic plasticity (Stokes et al. 1993). For instance, changes in age and size at first maturity can be caused by changes in growth rate (Reznick 1993,

Rijnsdorp 1993b). Since fishing typically is accompanied by a decrease in stock biomass, it may weaken density-dependent effects and cause a change in the onset of maturity simply through its effect on growth rate (Law 2000). For sole, there is evidence for density-mediated compensatory growth, although other factors such as temperature and changes in food availability might also play a role (Rijnsdorp & Van Beek 1991).

The effect of phenotypic plasticity can be disentangled from genetic effects by the reaction norm approach. By definition, a reaction norm describes which phenotypes will be expressed by a genotype under a certain range of environmental conditions (Stearns & Koella 1986). The probabilistic maturation reaction norm (PMRN) has been developed recently to estimate the reaction norm for the onset of maturity conditional on age and size (Heino et al. 2002, Barot et al. 2004). Environmental variability is assumed to be reflected in differences in somatic growth, so that any

environmental factor that has its effect through growth can be disentangled. This growth conditional PMRN has been applied to several stocks supporting a fisheries-induced evolutionary change in the onset of maturity (reviewed in Dieckmann & Heino 2007). Although the method has been criticized, as it does not disentangle genetic effects from environmental factors affecting maturation other than through size at age (Kraak 2007, Marshall & McAdam 2007), it can easily be extended to incorporate other variables as a third or higher dimension (Grift et al. 2007).

In this paper we investigate the extent to which the change in maturation in North Sea sole can be attributed to (1) phenotypic plasticity in response to the observed variations in growth, or (2) fisheries-induced evolutionary change using the PMRN approach. As environmental variation may affect maturation through more than variations in somatic growth, we explore the possible influence of condition by including it as a third dimension in the reaction norm approach, and analyse the influence of temperature and population density.

MATERIALS AND METHODS

Distribution and exploitation. The common sole *Solea solea* is distributed from the northwest African coast and the Mediterranean to the Irish Sea through the English Channel and the North Sea up to Skagerrak and Kattegat (Rijnsdorp et al. 1992). Spawning of North Sea sole occurs in the southern North Sea, with local concentrations in the German Bight, along the Belgian Coast and in the Thames estuary and the Wash (Fig. 1). Young sole are distributed in coastal waters and migrate to deeper offshore waters when they grow older but return every year in spring for spawning. North Sea sole is mainly caught by the Dutch fleet, which takes about 75% of the total international landings (ICES 2006) in a mixed demersal fishery with a minimum mesh size of 80 mm. The exploitation of sole has increased substantially since the 1960s following the introduction of the beam trawl (De Veen 1976).

Data. Fish samples have been collected monthly since 1957 from commercial landings covering the distribution area of sole in the North Sea and from autumn surveys since 1970. Market samples of individual fish are stratified into 5 market size categories and

record date, position, length (mm), weight (g), sex, gonad weight (g), maturity stage and age (yr). Age is determined from the pattern of growth rings in the otolith taking 1 January as birth date. Only females from samples collected in the southeastern North Sea (Areas 1 to 5; Fig. 1) were selected, and the analysis was restricted to Ages 1 to 6 of the Cohorts 1960 to 2002. In total, 21 969 observations were used for the analysis in the first seasonal sampling window, and 12 808 in the second (see below, Table 1).

Maturation. Maturity is determined by macroscopic inspection, and 8 maturity stages are distinguished that represent the seasonal development of the immature and mature gonad (Table 2; De Veen 1970): the immature gonad is classified as Stages I to II; developing ovaries as Stages III to IV representing the vitellogenic phase; the spawning stage when hyaline eggs
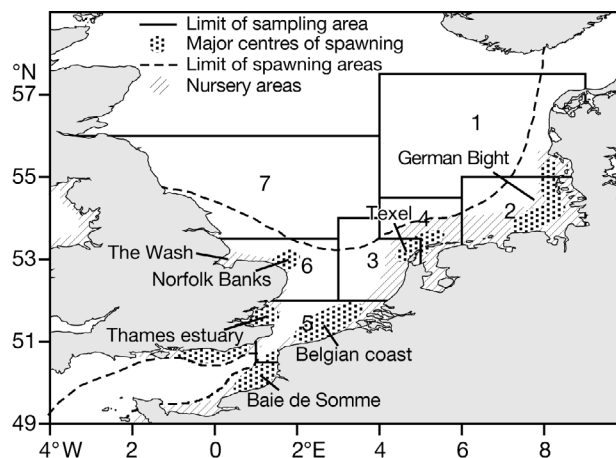

Fig. 1. Distribution of North Sea sole in sampling Areas 1 to 7. Spawning areas, indicated as the main centres of egg production, the limits of spawning activity and the nursery areas are based on Rijnsdorp et al. (1992)

Table 1. Sampling information for the data sets of female sole used a moving time window around the peak of spawning (2 mo before and 1 mo after the estimated maximum relative GSI) and a fixed time window (Nov, Dec and Jan)

	Age 1	Age 2	Age 3	Age 4	Age 5	Age 6	Total
Moving time window							
Total number	139	1226	7079	6490	4117	2918	21 969
Areas 3 & 4	131	1004	5195	4560	2885	2094	15 869
Areas 1 & 2	8	222	1884	1930	1232	814	6100
Average number per cohort	3	29	165	151	96	68	
Average percentage mature	3	10	78	97	98	99	
Fixed time window							
Total number	299	3870	3783	2531	1629	696	12 808
Areas 3–5	280	2908	2711	1788	1072	466	9225
Areas 1 & 2	19	962	1072	743	557	230	3583
Average number per cohort	0	7	90	88	59	38	
Average percentage mature	0	14	66	92	96	94	

Table 2. Macroscopic maturity stages of female sole. Stages I and II are interpreted as juvenile stages (–). Stages III to VIII are considered adult stages (+). After spawning is completed, the adult sole cycles from maturity Stage VIII to maturity Stage II into the growing phase of the following season

Stage	Interpretation	Observation
I	Juvenile (–)	Transparent and homogenous ovary, tight walls and small lumen, eggs indistinguishable
II	Resting (–)	Reddish translucent ovary, tight walls, eggs are distinguishable, lumen is filled with liquid
III	Vitellogenic (+)	Reddish gray to dark orange mostly opaque ovary, bigger and less tight than in II, rich in blood vessels, vitello-genesis (VTG) has started and some eggs contain yolk, lumen is big
IV	Vitellogenic (+)	Orange to reddish white, completely opaque and stiff ovary at half of its definite size, eggs are polygonal and all contain yolk
V	Ripe (+)	Orange to reddish white opaque ovary in its definite size, lumen is squeezed, eggs are round
VI	Spawning (+)	Mostly grayish-red translucent hyaline but partly still opaque ovary, resilient when pressed together, lumen contains spawn
VII	Half spent (+)	Gray to dark red completely translucent hyaline ovary, walls are very slack and bloody, lumen is big and filled with spawn and liquid
VIII	Spent (+)	Dark red translucent ovary, walls are very slack, lumen is very big mostly filled with liquid, very similar to II or in a transition towards II

are visible as Stages V to VI; and spent fish as stages VII to VIII. The start of vitellogenesis (VTG) ranges in North Sea sole from July to December (Fig. 2), varying between years, areas and also among individuals (Ramsay & Witthames 1996). Some fish might undergo abortive maturation, meaning that VTG has started but ovaries are not developed further than Stages III or IV (De Veen 1970, Ramsay & Witthames 1996).
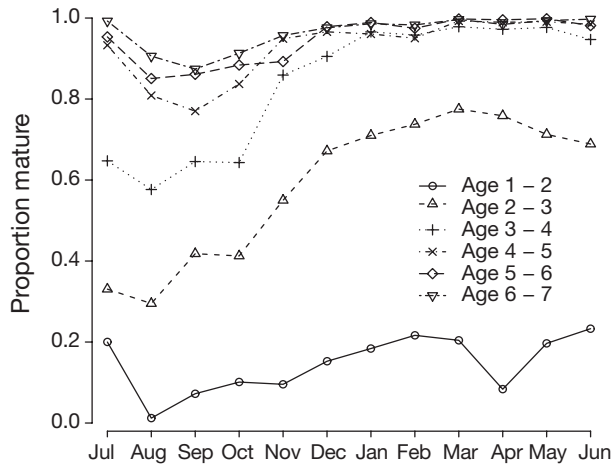

Fig. 2. Seasonal development of the average proportion of mature females by age-group in market samples between July in one year and June in the following year

Moving observation window. The ovary of a spent female (Stages VII to VIII) may look quite similar to the immature ovary (Stages I to II). Hence, the distinction between immature and mature females is more accurate in the time period when all mature females have started VTG or are in spawning stage (Stages III to VI), while none of them is spent. Consequently, the proportion of mature females is most reliably estimated in the period of peak spawning (Fig. 2; De Veen 1976, Ramsay & Witthames 1996). For an unbiased estimate of the proportion of mature fish, a moving time window around the peak of spawning was used to select data (Table 1). Because the timing of the spawning season varies annually in response to ambient water temperature (Van der Land 1991), peak spawning was determined as the date of the maximum ovary weight and was estimated for each year using the model

$$\log(o) = \beta_0 + \beta_1 \log(l) + \beta_2 d + \beta_3 d^2 + \beta_4 d^3 + \varepsilon \quad (1)$$

where o is the ovary weight, l is fish length and d is the day in the year on which the individual fish was caught. β represents the regression parameters and ε a normal error term (applies to all following equations).

Fixed observation window. The effects of weight and condition (i.e. relative weight: weight × length$^{-3.18}$) on maturation should ideally be analysed at the start of VTG (Wright 2007) because fish improve their condition as a consequence of readying to spawn in the subsequent season (Rijnsdorp 1990). However, because of the ambiguity in detecting maturation macroscopically at the start of VTG, a fixed window comprising the months November, December and January was used as a compromise between the moment where VTG starts and the moment where detection of maturity becomes reliable (Table 1, Fig. 2).

Probabilistic maturation reaction norm. The PMRN is defined as the probability that an immature individual becomes mature at a particular age and size (Heino et al. 2002) and thereby overcomes the confounding effects of growth and mortality. For sole, first time and repeat spawners cannot be distinguished, therefore the probability of becoming mature (p) was estimated by a refinement of the PMRN method (Barot et al. 2004) and is estimated on a yearly basis as:

$$p(a,s,x) = \frac{m(a,s,x) - m(a-1,s-\Delta s,x-\Delta x)}{1 - m(a-1,s-\Delta s,x-\Delta x)} \quad (2)$$

where m is the probability of being mature, a is age, s is size, x is any other additional factor possibly affecting maturation (e.g. condition, see Grift et al. 2007), Δs is the age specific growth increment (in length or weight) and Δx is the age specific change in x. Eq. (2) relies on the simplifying assumption that the growth and mortality rates are the same for immature and mature individuals. Although this is not expected to apply accurately, Barot et al. (2004) confirmed that the method is robust to the relaxation of this assumption. The probability of being mature $m(a,s,x)$ was estimated from individual data by logistic regression. Model selection was done by forward selection of variables based on the Akaike information criterion (AIC) allowing for all possible 1- and 2-way interactions. Two-way interactions were considered when both components of the interaction had been selected:

$$\text{logit}(m) = \beta_0 + \beta_c c + \beta_a a + \beta_l l + \beta_{ca} c \times a + \beta_{cl} c \times l + \beta_{al} a \times l + \varepsilon \quad (3)$$

$$\text{logit}(m) = \beta_0 + \beta_c c + \beta_a a + \beta_w w + \beta_{ca} c \times a + \beta_{cw} c \times w + \beta_{aw} a \times w + \varepsilon \quad (4)$$

$$\text{logit}(m) = \beta_0 + \beta_c c + \beta_a a + \beta_l l + \beta_k k + \beta_{ca} c \times a + \beta_{cl} c \times l + \beta_{al} a \times l + \beta_{lk} l \times k + \varepsilon \quad (5)$$

where m is the probability of being mature, c is cohort as a categorical variable, a is age, l is length, w is weight and k is condition. Because of differences in the interpretation of the timing of effects of length, weight and condition on maturation (see 'Moving observation window' and 'Fixed observation window' above) Model 3 (Eq. 3) was applied to the moving window, and Models 4 (Eq. 4) and 5 (Eq. 5) to the fixed window data set (Table 1).

Growth increments. Annual length increments were estimated by fitting Von Bertalanffy growth curves per cohort. Curves were fitted to mean length at age data, after accounting for length stratification of market samples (Age > 3), survey and backcalculation data (Age ≤ 3). For age groups up to Age 3, the mean length from market samples will be biased because small individuals may escape through the mesh or may be discarded if below minimum marketable size. Therefore, pre-recruit survey data and back-calculation data of Millner & Whiting (1996) were used (see Fig. 3a). The mean growth increments did not differ substantially between the German Bight and the Dutch-Belgian Coast spawning stocks ($r = 0.98$, $R^2 = 0.96$).

Weight and condition increments per cohort were estimated as follows. First, the length–weight relationship was modeled assuming a fixed seasonal pattern:

$$\log(w) = \beta_0 + \beta_1 \log(l) + \beta_2 \text{yr} + \beta_3 d + \beta_4 d^2 + \beta_5 d^3 + \varepsilon \quad (6)$$

where d is the day in year and year (yr) is a factor. The length–weight scaling factor was $\beta_1 = 3.18$ and has been used throughout the analysis to define the condition factor as $k = 100\, w\, l^{-3.18}$. The mean weight at age per cohort was then calculated from the mean length and corrected for the bias introduced when estimating mean weight from mean length. Given the observed coefficient of variation in length at age, a correction factor was estimated from an empirically derived relationship between the bias and the coefficient of varia-
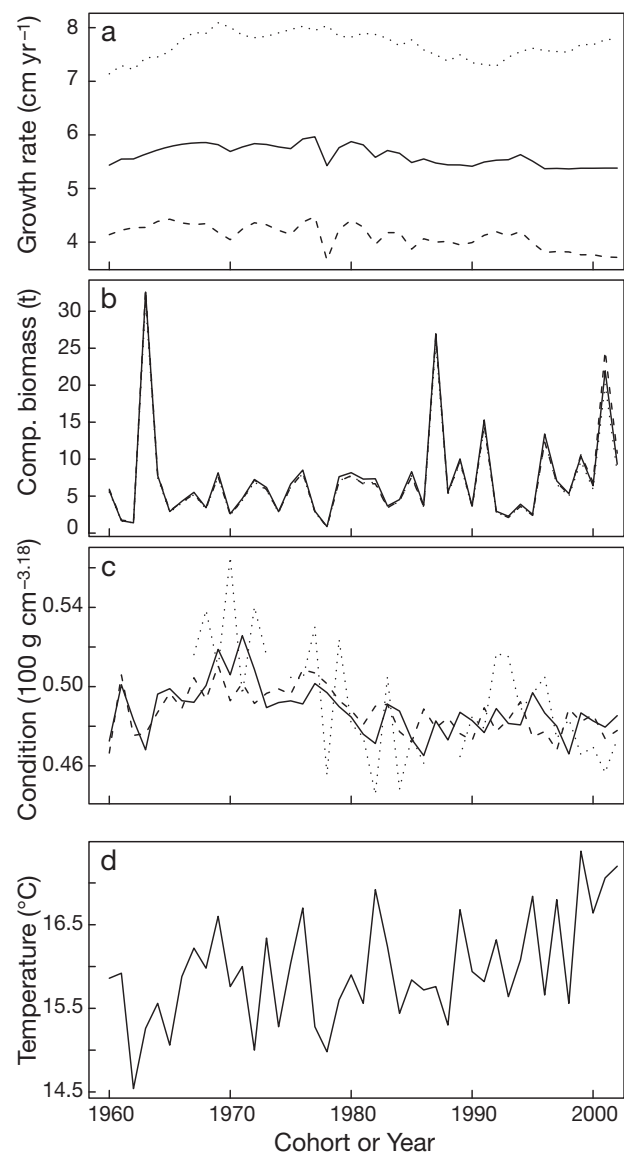

Fig. 3. Time trends per cohort in (a) growth rate as yearly length increments; (b) individually experienced competitive biomass; and (c) condition. (d) Temperature per year. Panels (a) to (c): ··· Age 2; — Age 3; --- Age 4

tion assuming a normal distribution. Because growth increments depend on size rather than age, the annual increments in length, Δl, and weight, Δw, (see Δs in Eq. 2) realized when reaching age a, were estimated separately for each cohort as

$$\Delta l_a = l_a - l_{a-1} = \beta_0 + \beta_1 \, l_a + \varepsilon \qquad (7)$$

$$\Delta w_a = w_a - w_{a-1} = \beta_0 + \beta_1 w_a + \beta_2 w_a^{2/3.18} + \varepsilon \qquad (8)$$

The so-modeled increments are in line with the theoretical expectation given Von Bertalanffy growth and the length–weight scaling used here.

Diagnostics. To visualize the reaction norm and its change over time we use the reaction norm midpoints in length (Lp50) and weight (Wp50) at which the probability of becoming mature is 0.5. Reaction norm midpoints were obtained by linear interpolation. In the case of the 3-dimensional reaction norm, the midpoints were estimated for a fixed level of the third dimension (condition). Because the reaction norm midpoints are obtained after various steps of calculation, it is not possible to obtain confidence intervals directly and, therefore, they were bootstrapped by resampling the individuals in the original data set with replacement 1000 times, for each cohort. To test for a trend in the Lp50 and Wp50 at a given age, the reaction norm midpoints were regressed to cohort as a variable, weighted by the inverse of the bootstrap variance of each midpoint. Confidence intervals were estimated from the 95% percentile in the bootstrap replicates. Because the onset of maturation might differ between subpopulations, the analysis was run separately for the spawning subpopulation of the German Bight (sampling Areas 1 and 2; Fig. 1) and the Dutch and Belgian coast (Areas 3, 4 and 5; Fig. 1) as well as for the pooled data of the 2 stocks. Because the results of the subpopulations were similar, only results for the pooled data set representing the whole southern North Sea are presented here. Because 95% of the sole matured at Ages 2, 3 and 4 (see Table 1) we only consider these ages for diagnostics.

Alternative factors. To evaluate alternative variables, which could not be incorporated on the individual level PMRN methodology, the Lp50 and Wp50 estimates were regressed against annual estimates of condition, temperature and competitive biomass (Fig. 3). Condition was estimated as the average condition in the third and fourth quarters. Temperature was measured at a coastal station at 53° N, 5° E and averaged from June to October (Van Aken 2003). Competitive biomass, i.e. the biomass with which an individual fish has to compete for resources, was estimated per age group from virtual population analysis stock numbers and weights (ICES 2006) and indices of mean crowding following Rijnsdorp & Van Beek (1991).

Combinations of alternative factors were regressed for age class (a) to explain the estimated reaction norm midpoints over cohorts, weighted by the inverse of the corresponding bootstrap variances:

$$Sp50_{c,a} \sim cohort + \Sigma_\Delta T_{c+a-\Delta} + \Sigma_\Delta K_{c,a-\Delta} + \Sigma_\Delta B_{c,a-\Delta} + \varepsilon_{c,a} \qquad (9)$$

where $Sp50$ is either Lp50 or Wp50, *cohort* is cohort as a variable, $T_{c+a-\Delta}$ is the average temperature in year $c + a - \Delta$, $K_{c,a-\Delta}$ is the average condition in cohort c at age $a - \Delta$, and $B_{c,a-\Delta}$ is the estimated competitive biomass in cohort c at age $a - \Delta$. Σ_Δ denotes the time lagged variables with a lag Δ being an element of $\{1, 2, \ldots, a\}$. Model selection was based on the smallest prediction error and the number of parameters was increased forwards-wise. For a given number of parameters, the combination of time lagged variables giving the lowest predicted mean squared error (MSE) in a leave-one-out cross-validation is considered to be the most meaningful model.

RESULTS

Model properties

Cohort, age and size (length or weight), as well as all the 2-way interactions, significantly affected the probability of being mature (Table 3). The length-based Model 3, using a moving time window, gave a better fit and prediction than the weight- and condition-based models (Models 4 & 5) that used a fixed time window. Model 3 explained about 14% more of the null deviance than the fixed window models (Table 3). Considering length and condition as explanatory variables (Model 5) instead of weight (Model 4) results in a lower AIC (Table 3). All models had similarly high classification performance. The sensitivity, i.e. the proportion of correctly classified individuals among matures, was high, but specificity, i.e. the proportion of correctly classified individuals among immatures, was relatively low (Table 3).

Length- and weight-based reaction norms

The estimated reaction norms are close to linear and tend to have negative slopes, suggesting that the size at which sole attains a certain probability to mature decreases with age. This age effect is weaker in the weight reaction norm, such that the probability of becoming mature is determined by weight rather than age. The width of the reaction norm is narrower for length than for weight. The distance between Lp10 and Lp90 is on average 23% of the Lp50

Table 3. Model properties: Akaike information criterion (AIC), proportion of null deviance explained by the model, proportion of correctly classified observations, sensitivity (proportion of correctly classified positives, i.e. matures) and specificity (proportion of correctly classified negatives, i.e. immatures) of Models 3, 4 and 5. Model 3 was built on data from the moving window data selection, Models 4 and 5 were built on data from the fixed window data selection. AIC should not be compared between models obtained from different data sets

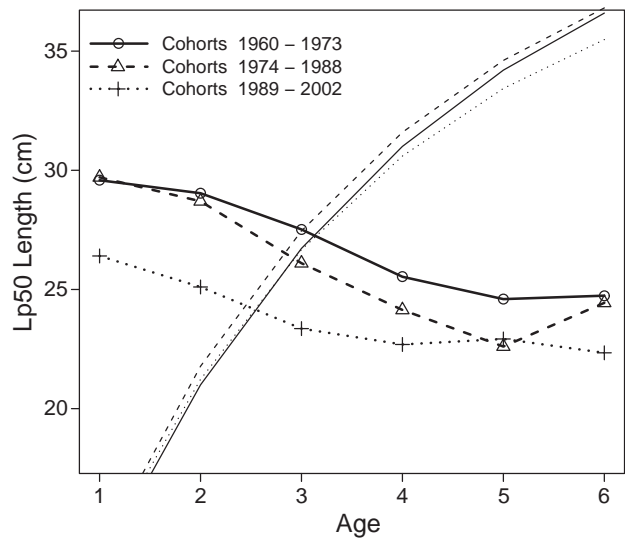| | Moving window | Fixed window | |
	Model 3	Model 4	Model 5
AIC	8425	7086	7015
Explained deviance	0.54	0.40	0.41
Correct classification	0.92	0.89	0.89
Sensitivity	0.98	0.97	0.97
Specificity	0.59	0.49	0.51

Fig. 4. Reaction norms for Cohorts 1960 to 1973 (—), 1974 to 1988 (---) and 1989 to 2002 (···) estimated as the average Lp50. Thin lines without symbols show mean length at age (cm) of the corresponding cohorts

(Model 3), whereas the distance between Wp10 and Wp90 is on average 65% of the Wp50 (Model 4). Over time, the reaction norms shifted towards younger ages and smaller sizes, with the largest change occurring in the 1980s (Figs. 4 & 5). The intersection of average growth with the reaction norm shows that the age and length at which 50% of the females became mature shifted from about 3.1 yr and 27.3 cm in the 1960s to about 2.5 yr and 24.1 cm in the 1990s (Fig. 4). The Lp50 and Wp50, reflecting the intercepts of the reaction norm, have typically decreased significantly over time at all ages (Fig. 5). The decrease in the Lp50 and Wp50, weighted by the inverse bootstrap variances, is significant for all ages ($p < 10^{-4}$ for all ages in Fig. 5). Inter-annual variation was too large to detect a significant trend in the reaction norm slopes (data not shown). Since most individuals become mature at Age 3 (Table 1), the method is most precise for this age group. For 3 yr old females, the weighted regressions predict that the size at which 50% become mature decreased from 28.6 cm (251 g) in the 1960 cohort to 24.6 cm (128 g) in the 2002 cohort ($p < 10^{-5}$). This trend corresponds to a change of -0.10 cm yr^{-1} in the Lp50 and -3 g yr^{-1} in the Wp50. For the Lp50 values, the minimal marketable size of 24 cm seems to operate as an attractor: the decline of the midpoints ends with an Lp50 of about 24 cm at all ages.

Three-dimensional reaction norm

Condition in the 3-dimensional reaction norm (Model 5) is positively associated with the probability of becoming mature. The Lp50 of the reaction norms, estimated for all values within the 95% CI of observed condition (0.408 to 0.604), indicated that a change of

10% in condition results in a change of about 1 cm in the Lp50. The decrease in Lp50 over time for any given value of condition factor, and, thus, the changes in condition observed in the study period, could not explain the observed decrease in the 2-dimensional Lp50s. The weighted regression on midpoints of the 3-dimensional reaction norm, bootstrapped for the median condition of $k = 0.5$, predicts a decrease in Lp50 from 28.9 cm in 1960 to 23.2 cm in 2002. The decline of -0.14 cm yr^{-1} is stronger than in the 2-dimensional reaction norm. The reaction norm width (21% of the midpoints for condition of $k = 0.5$, Model 5) is similar to the width in the length reaction norm (23% of the midpoints, Model 3) and much smaller than the width in the weight reaction norm (65% of the midpoints, Model 4).

Alternative factors

The alternative factors analysed for their effect on the maturation all showed significant trends over the study period (Fig. 3). Growth rates have decreased significantly (Age 2: -0.045 cm yr^{-2}; Age 3: -0.032 cm yr^{-2}; Age 4: -0.022 cm yr^{-2}; all $p < 10^{-7}$) as does condition (Age 2: -0.13 g $cm^{-3.18}$ yr^{-1}; Ages 3 & 4: -0.04 g $cm^{-3.18}$ yr^{-1}; all $p < 10^{-2}$). Temperature increased significantly (0.03°C yr^{-1}, $p < 10^{-3}$), while competitive biomass showed a peak in the early 1960s, and also around 1990 and 2000. Table 4 presents the variables for the time series in Model 9 (Eq. 9) that give the lowest predicted MSE in a leave-one-out cross-validation for Age 3. The general
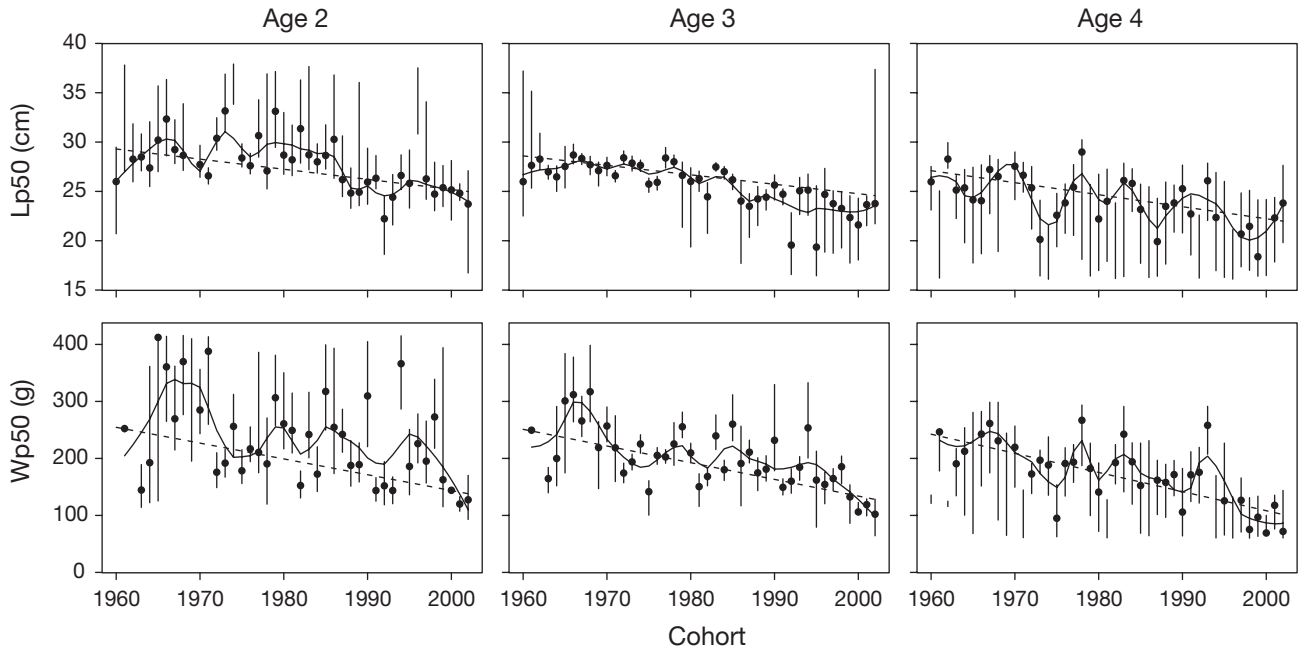
Fig. 5. Reaction norm midpoints Lp50 and Wp50 over time (●), bootstrapped 95% percentiles (vertical bars), trend regression weighted by the inverse bootstrap variances (---) and fit with a non-parametric smoother (—). All trends are significant on a level of $\alpha = 10^{-4}$

pattern of variable selection is consistent for Lp50 and Wp50 for all ages and for all time lags Δ: cohort is always selected as the first variable, while the variables to be selected next are temperature in the year of birth, if this year is included in the analysis, competitive biomass otherwise (Table 4). The MSE only weakly improves by further introduction of variables after cohort, and would increase again by introducing more variables than those displayed in Table 4 (results not shown).

DISCUSSION

Evidence for evolutionary change

This study shows that the PMRN in North Sea sole shifted downwards since the 1960 cohort, which is in agreement with the expectation of a fisheries-induced evolutionary change. In an evolutionary fitness optimization framework, the reaction norm intercept is predicted to decrease and the slope is predicted to increase slightly if exposed to size-selective harvesting

Table 4. Selected models to explain the variation in the Lp50 and Wp50 of Age 3, alternative variables based on the lowest predicted mean square error (MSE) in a leave-one-out cross-validation by allowing the effectors to act on a particular age (Δ fixed) or on all ages together before maturation (Δ element of {1, 2, 3}). Best models were determined by comparing the MSEs from all possible combinations of variables for a given number of variables. Only models for which the cross-validated MSE is lower than it would be if a further variable was introduced are displayed. *: significant at $\alpha = 0.05$. T, B, K: See Eq. 9

Impact period	Number of variables	Selected variables in model	Signs of parameters	MSE
Lp50 at Age 3				
–	0			5.278
Age 0, $\Delta = 3$	1	Cohort*	–	2.320
	2	Cohort*+$T_{age\,0}$*	– –	2.172
Age 1, $\Delta = 2$	1	Cohort*	–	2.320
Age 2, $\Delta = 1$	1	Cohort*	–	2.320
All ages,	1	Cohort*	–	2.320
$\Delta = \{1, 2, 3\}$	2	Cohort*+$T_{age\,0}$*	– –	2.172
Wp50 at Age 3				
–	0			2818
Age 0, $\Delta = 3$	1	Cohort*	–	1816
Age 1, $\Delta = 2$	1	Cohort*	–	1816
	2	Cohort*+ B_{age1}*	– –	1719
	3	Cohort*+ B_{age1}*+ K_{age1}	– – +	1452
Age 2, $\Delta = 1$	1	Cohort*	–	1816
	2	Cohort*+ B_{age2}*	– –	1709
All ages,	1	Cohort*	–	1816
$\Delta = \{1, 2, 3\}$	2	Cohort*+ B_{age2}*	– –	1709
	3	Cohort*+ B_{age2}*+ K_{age1}	– – +	1448
	4	Cohort*+ B_{age1}*+ K_{age1} + K_{age2}	– – + –	1430

(Ernande et al. 2004). The changes in the reaction norm intercepts found here are consistent with this theoretical expectation (see Fig. 5), but noise in our data was too high to detect a trend in the slopes. Based on the assumption that environmental variation affecting maturation is reflected in variations in the growth rate (Heino et al. 2002), a genetic change in maturation can thus be disentangled from the phenotypic plasticity in maturation. The length-based reaction norm is more reliably estimated than the weight-based reaction norm because of the ambiguity in data selection and corresponding interpretation for the weights (see below) resulting in a smaller data set and higher uncertainty in the reaction norm estimates. However, other factors in addition to size, such as body condition (Kjesbu et al. 1991, Trippel & Neil 2004) or temperature (Dhillon & Fox 2004, Kraak 2007), may affect the probability of becoming mature.

Including condition in the analysis as a third dimension revealed that a higher condition is associated with lower Lp50 values, thus fish in good condition have a higher probability of becoming mature than fish of the same age and length in poor condition. The same effect has been suggested in numerous other species (Kjesbu et al. 1991, Trippel & Neil 2004) and is likely to be (partly) related to the storage of energy during the feeding period that can be metabolized later for reproduction (Rijnsdorp 1990). Although condition contributed significantly to variation in maturation, it could not explain the trends in the midpoints of the 2-dimensional maturation reaction norm over time: the decrease in Lp50 at any given value of condition (see Fig. 6) indicates that the earlier maturation is not due to a time trend in the condition factor. Given the positive association between condition and the probability to become mature, the pattern of observed conditions (Fig. 3c) suggests that by not taking condition into account, the midpoints of the 2-dimensional reaction norm might have been underestimated in the first decades, when condition was high, and overestimated in the last decades when condition was low. Indeed, the estimated shift of the midpoints in the 3-dimensional reaction norm (–0.14 cm yr^{-1}, Fig. 6) is stronger than in the 2-dimensional reaction norm not accounting for condition (–0.10 cm yr^{-1}, Fig 5). The effect of condition in sole seems to be very weakly age-dependent, but it is changing as the maturation reaction norm shifts downwards. Hence, for an early maturing fish, it is getting more important to be in good condition. However, the interpretation of the results in terms of the role of condition in the maturation process is not straightforward (Wright 2007). The role of condition was tested in winter when all maturing females have started VTG, several months after the maturation 'decision' was made. Although we used a fixed win-
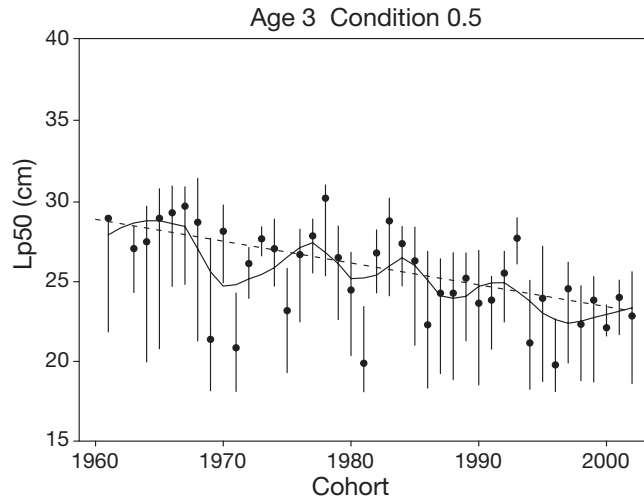

Fig. 6. Reaction norm Lp50 midpoints estimated with Model 5 for Age 3 and condition 0.5 (●), bootstrap 95% percentiles (vertical bars), trend regression weighted by the inverse bootstrap variances (- - -), fit with a non-parametric smoother (—). The trend is significant on a level of $\alpha = 10^{-5}$

dow as close as possible to the start of VTG, the positive effect of weight and condition on the probability to become mature is not necessarily affecting the 'decision' to become mature, but is more likely a consequence of it, because fish seem to improve their condition as a consequence of the 'decision' to become mature in the subsequent season (Rijnsdorp 1990). In addition to the different proportions of maturity at age between the 2 data sets, this may also explain the better fit and performance of Model 3 in comparison to Models 4 and 5 (Table 3). Also, the higher uncertainty of the Wp50 estimates might be at least partly due to inter-annual differences in the seasonal ovary development. Unfortunately, this ambiguity has no practical solution.

Rate of evolutionary response

Because of the difficulty of measuring selection differentials (Sinclair et al. 2002) and heritabilities, especially in the wild, the rate of evolution is not well understood. Experiments have shown that a change in the pattern and rate of size-selective mortality caused major evolution of life history characteristics in 18 generations of guppies *Poecilia reticulate* (D. A. Reznick et al. 1990) and the growth rate of Atlantic silverside *Menidia menidia* evolved within just 4 generations (Conover & Munch 2002). Genetic analyses of rapid life history changes following introduction to novel environments are also well established in several species in the wild (D. N. Reznick et al. 1997, Quinn et al. 2001), and hence, the hypothesis of a rapid

fisheries-induced evolution in sole is feasible. The rate of change in the length maturation reaction norm (Model 3) does not seem to be linear, as stepwise change in its reaction norm midpoints seems to occur in the 1980s (Fig. 5). Possible reasons are that: (1) the pattern is due to a change in condition; (2) there was still a considerable proportion of old individuals in the population that contributed to the next generations, reducing the evolutionary response; (3) the selection differential increased after the 1970s; or (4) gene flow between subpopulations affected the changes in reaction norms. Explanation (1) is plausible since the trend in midpoints of reaction norms including weight or condition does not show the same stepwise pattern (Figs. 5 & 6). Explanation (2) is likely since larger fish have a relatively higher fecundity and may contribute disproportionately to reproductive success (Trippel & Neil 2004). Explanation (3) is likely since fishing mortality has increased over time. Since the increase in fishing mortality is not synchronized between subpopulations (ICES 2006), a delay in the evolutionary response due to gene flow (Explanation [4]) may also have played a role (see Conover et al. 2005). The downward trend in the reaction norm is expected to stop, once it falls below the size threshold where selection no longer has any effect (Ernande et al. 2004), e.g. below the minimum size of capture that is close to the minimum landing size of 24 cm. In recent years, the Lp50 in all age groups came close to this threshold and therefore the trend might weaken in the following years.

Alternative factors

It is known that the onset of maturation may be affected by condition (Kjesbu et al. 1991, Trippel & Neil 2004) or other factors such as temperature (Dhillon & Fox 2004, Kraak 2007), in the juvenile period 1 or more years prior to first maturity. The influences of average temperature, condition and competitive biomass were analysed on the population level in the years before maturation. The results show clearly that none of these possible factors could explain the variance in the midpoints better than the trend variable *cohort* (Table 4). Hence, the evolutionary interpretation is the most plausible explanation of the observed shifts in the maturation reaction norm. On top of this trend, temperature in the year of birth and competitive biomass contribute to further reduce the prediction error. Condition, having a considerable effect if introduced as a third dimension in the reaction norm, does not seem to play an important role in the population level, because condition might be much more affected by individual behavior than, for example, either temperature or experienced competitive biomass.

Temperature is known to affect early life history stages, as it may affect age at maturity through its effect on juvenile growth rates (Charnov & Gillooly 2004). However, growth-independent effects of temperature on maturation have also recently been suggested (Dhillon & Fox 2004). Temperature can accelerate developmental rates other than growth (Baynes & Howell 1996, Fuiman et al. 1998) and, therefore, the onset of maturation might be affected by temperature in early life. Since the temperature in the year of birth acting on the larval stage, metamorphosis and early juvenile stages explains most of the variance in addition to the intrinsic trend for the Lp50 (Table 4), our results support such an early life history stage effect.

Similarly, competitive biomass likely has its main effect on maturation through density-dependent growth. However, the significance of competitive biomass in explaining variance in the weight reaction norm midpoints in addition to the intrinsic trend suggest that it may also have a growth-independent effect, driving the fish to accept an earlier maturation at a smaller size if population densities are high (sign of parameter is negative, Table 4). Maturation may be influenced by social factors, e.g. the probability to encounter conspecifics, as shown in the coral-dwelling fish *Gobiodon erythrospilus* (Hobbs et al. 2004). This mechanism might explain a negative correlation of competitive biomass with the reaction norm midpoints (Table 4). By contrast, the presence of dominant adults might suppress maturation of subordinate juveniles as shown in males of the freshwater platyfish genus *Xiphophorus* (Borowsky 1978), hermaphrodites of the coral reef anemone fish genus *Amphiprion* (Fricke & Fricke 1977) and females of the temperate wrasse *Pseudolabrus celidotus* (Jones & Thompson 1980). This would result in a positive correlation of competitive biomass with the reaction norm midpoints.

Another possible cause for changes in the reproductive cycle is that fish might accumulate increased levels of vitellogenin in their tissue due to endocrine disrupting agents present in sewage effluents (Oberdörster & Cheek 2001). Estrogenic pollution (oestradiol or analogues) was found to be associated with elevated concentrations of vitellogenin in flatfish species in the Dutch Wadden Sea (Janssen et al. 1997) and also in samples taken offshore at the Dogger Bank in the North Sea, leading to intersex (Stentiford & Feist 2005). However, since the reproductive effects of increased vitellogenin concentrations in fish are often associated with reproductive impairment and reduced fecundity (Oberdörster & Cheek 2001), it is not clear if increased vitellogenin would accelerate or delay maturation or how these effects would become manifest on the pop-

ulation level. Nevertheless, increased levels of vitellogenin in female *Solea solea* could lead to misclassification of immatures as matures in the macroscopic maturity staging. Therefore, exposure to high concentrations of contaminants during the early juvenile phase in coastal nursery grounds may have influenced the results presented here.

Sources of error

Our results may be influenced by the assumption made to estimate reaction norms: that mature and immature fish have similar growth and mortality rates. However, a sensitivity analysis has shown that the PMRN methodology is robust with respect to differences in growth rate (Barot et al. 2004) and furthermore, these biases would apply to all cohorts equally and thus would not qualitatively change the trends over time. Macroscopic staging of maturity could be erroneous by misclassifying spent or unripe females as juveniles. However, by selecting data around the peak spawning period and applying a moving time window, this error will be negligible. Abortive maturation, in the sense that fish develop until maturity Stage III but not further (De Veen 1970, Ramsay & Witthames 1996), could have biased the PMRN. However, considering maturity Stages I, II and III as juveniles instead of only Stages I and II changed neither our results (not shown) nor our conclusions.

Implications for management

The evidence for fisheries-induced evolution is strong and persistent. Fisheries-induced evolution is likely to decrease maximum sustainable yield (Law & Grey 1989, Heino 1998) and finding practical strategies to reverse the decreasing trend in age and size at maturation will be increasingly difficult. Recovery from a genetic change will be much slower than could be expected from just lower abundance, since selection pressure toward the original genotype in the absence of fishing may be much weaker than selection caused by intensive fishing (Law 2000). If directional selection is strong, the original genotype may have been lost, and surviving genotypes might have a lower fitness in an environment with a lower exploitation rate (Conover et al. 2005). Advanced maturation will result in a reduced size at age of mature fish and thus in a lower biomass per age group (Heino 1998). Whether the reduced growth rates observed in North Sea sole (Fig. 3a) are a direct consequence of the observed shift in the onset of maturation or due to a general deterioration of the growing conditions needs to be analysed

in more detail. A direct selection on growth rates, however, seems less likely because in iteroparous species with indeterminate growth, the improved survival through slower growth is unlikely to offset the loss in fecundity (Heino & Kaitala 1999). Even if earlier maturation will positively affect the spawning stock biomass, this may not necessarily translate to an increase in recruitment, since fecundity and egg viability are positively correlated with maternal size (Trippel & Neil 2004), and therefore earlier maturing, smaller individuals will contribute less to reproduction.

LITERATURE CITED

Barot S, Heino M, O'Brien L, Dieckmann U (2004) Estimating reaction norms for age and size at maturation when age at first reproduction is unknown. Evol Ecol Res 6:659–678

Baynes SM, Howell BR (1996) The influence of egg size and incubation temperature on the condition of *Solea solea* (L) larvae at hatching and first feeding. J Exp Mar Biol Ecol 199:59–77

Borowsky R (1978) Social inhibition of maturation in natural populations of *Xiphophorus variatus* (Pisces Poeciliidae). Science 201:933–935

Charnov EL, Gillooly JF (2004) Size and temperature in the evolution of fish life histories. Integr Comp Biol 44:494–497

Conover DO, Munch SB (2002) Sustaining fisheries yields over evolutionary time scales. Science 297:94–96

Conover DO, Arnott SA, Walsh MR, Munch SB (2005) Darwinian fishery science: lessons from the Atlantic silverside (*Menidia menidia*). Can J Fish Aquat Sci 62:730–737

De Veen JF (1970) On some aspects of maturation in the common sole *Solea solea* (L.). Ber Dtsch Wiss Komm Meeresforsch 21:78–91

De Veen JF (1976) Changes in some biological parameters in North-Sea sole (*Solea solea* L). J Conseil 37:60–90

Dhillon RS, Fox MG (2004) Growth-independent effects of temperature on age and size at maturity in Japanese medaka (*Oryzias latipes*). Copeia 2004:37–45

Dieckmann U, Heino M (2007) Probabilistic maturation reaction norms: their history, strengths, and limitations. Mar Ecol Prog Ser 335:253–269

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

Fricke H, Fricke S (1977) Monogamy and sex change by aggressive dominance in coral reef fish. Nature 266:830–832

Fuiman LA, Poling KR, Higgs DM (1998) Quantifying developmental progress for comparative studies of larval fishes. Copeia 1998:602–611

Grift RE, Heino M, Rijnsdorp AD, Kraak SBM, Dieckmann U

(2007) Three-dimensional maturation reaction norms for North Sea plaice. Mar Ecol Prog Ser 334:213–224

Heino M (1998) Management of evolving fish stocks. Can J Fish Aquat Sci 55:1971–1982

Heino M, Kaitala V (1999) Evolution of resource allocation between growth and reproduction in animals with indeterminate growth. J Evol Biol 12:423–429

Heino M, Dieckmann U, Godo OR (2002) Measuring probabilistic reaction norms for age and size at maturation. Evolution 56:669–678

Hobbs JPA, Munday PL, Jones GP (2004) Social induction of maturation and sex determination in a coral reef fish. Proc R Soc Lond Ser B Biol Sci 271:2109–2114

ICES (2006) Report of the working group on the assessment of demersal stocks in the North Sea and Skagerrak. ICES CM 2006/ACFM:09

Janssen PAH, Lambert JGD, Vethaak A, Goos HJT (1997) Environmental pollution caused elevated concentrations of oestradiol and vitellogenin in the female flounder, *Platichthys flesus* (L). Aquat Toxicol 39:195–214

Jones GP, Thompson SM (1980) Social inhibition of maturation in females of the temperate wrasse *Pseudolabrus celidotus* and a comparison with the blennioid *Tripterygion varium*. Mar Biol 59:247–256

Jørgensen T (1990) Long-term changes in age at sexual maturity of Northeast Arctic cod (*Gadus morhua* L). J Conseil 46:235–248

Kjesbu OS, Klungsoyr J, Kryvi H, Witthames PR, Walker MG (1991) Fecundity, atresia, and egg size of captive Atlantic cod (*Gadus morhua*) in relation to proximate body composition. Can J Fish Aquat Sci 48:2333–2343

Kraak SBM (2007) Does the probabilistic maturation reaction norm approach disentangle phenotypic plasticity from genetic change? Mar Ecol Prog Ser 335:295–300

Law R (2000) Fishing, selection, and phenotypic evolution. ICES J Mar Sci 57:659–668

Law R, Grey DR (1989) Evolution of yields from populations with age-specific cropping. Evol Ecol 3:343–359

Marshall CT, McAdam BJ (2007) Integrated perspectives on genetic and environmental effects on maturation can reduce potential for errors of inference. Mar Ecol Prog Ser 335:301–310

Millner RS, Whiting CL (1996) Long-term changes in growth and population abundance of sole in the North Sea from 1940 to the present. ICES J Mar Sci 53:1185–1195

Oberdörster E, Cheek AO (2001) Gender benders at the beach: endocrine disruption in marine and estuarine organisms. Environ Toxicol Chem 20:23–36

Quinn TP, Kinnison MT, Unwin MJ (2001) Evolution of chinook salmon (*Oncorhynchus tshawytscha*) populations in New Zealand: pattern, rate, and process. Genetica 112:493–513

Ramsay K, Witthames P (1996) Using oocyte size to assess seasonal ovarian development in *Solea solea* (L). J Sea Res 36:275–283

Reznick DA, Bryga H, Endler JA (1990) Experimentally induced life-history evolution in a natural population. Nature 346:357–359

Reznick DN (1993) Norms of reaction in fishes. In: Stokes TK, McGlade JM, Law R (eds) The exploitation of evolving resources. Springer-Verlag, Berlin

Reznick DN, Shaw FH, Rodd FH, Shaw RG (1997) Evaluation of the rate of evolution in natural populations of guppies (*Poecilia reticulata*). Science 275:1934–1937

Rijnsdorp AD (1990) The mechanism of energy allocation over reproduction and somatic growth in female North Sea plaice, *Pleuronectes platessa* L. Neth J Sea Res 25:279–289

Rijnsdorp AD (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 AD (1993b) Relationship between juvenile growth and the onset of sexual maturity of female North Sea plaice, *Pleuronectes platessa*. Can J Fish Aquat Sci 50:1617–1631

Rijnsdorp AD, Van Beek FA (1991) Changes in growth of plaice *Pleuronectes platessa* L and sole *Solea solea* (L) in the North Sea. Neth J Sea Res 27:441–457

Rijnsdorp AD, Van Beek FA, Flatman S, Millner R, Giret M, De Clerck R (1992) Recruitment of sole stocks, *Solea solea* (L.), in the northeast Atlantic. Neth J Sea Res 29(1–3):173–192

Sinclair AF, Swain DP, Hanson JM (2002) Measuring changes in the direction and magnitude of size-selective mortality in a commercial fish population. Can J Fish Aquat Sci 59:361–371

Stearns SC, Koella JC (1986) The evolution of phenotypic plasticity in life history traits: predictions of reaction norms for age and size at maturity. Evolution 40:893–913

Stentiford GD, Feist SW (2005) First reported cases of intersex (ovotestis) in the flatfish species dab *Limanda limanda*: Dogger Bank, North Sea. Mar Ecol Prog Ser 301:307–310

Stokes TK, McGlade JM, Law R (1993) The exploitation of evolving resources. Proc Intl Conf, Julich, Germany, 3–5 Sept 1991. Springer-Verlag, Telos

Trippel EA, Neil SRE (2004) Maternal and seasonal differences in egg sizes and spawning activity of northwest Atlantic haddock (*Melanogrammus aeglefinus*) in relation to body size and condition. Can J Fish Aquat Sci 61:2097–2110

Van Aken HM (2003) 140 years of daily observations in a tidal inlet (Marsdiep). ICES Marine Science Symposia 219: Hydrobiological variability in the ICES Area, 1990–1999, p 359–361

Van der Land MA (1991) Distribution of flatfish eggs in the 1989 egg surveys in the southeastern North Sea, and mortality of plaice and sole eggs. Neth J Sea Res 27:277–286

Wright PJ (2007) Understanding the maturation process for field investigations of fisheries-induced evolution. Mar Ecol Prog Ser 335:279–283