Brage IMR – 

Havforskningsinstituttets institusjonelle arkiv


Ved lenking til artikkelen skal det lenkes til post i Brage IMR, ikke direkte til pdf-fil.

Brage IMR –

Institutional repository of the Institute of Marine Research

This is the author’s last version of the article after peer review and is not the publisher’s version, usually referred to as postprint. You will find a link to the publisher’s version in Brage IMR. It is recommended that you obtain the references from the publisher’s site.

Linking to the article should be to the Brage-record, not directly to the pdf-file.
Investigating the drivers of maturation dynamics in Barents Sea haddock (*Melanogrammus aeglefinus*)

Jennifer A. Devine a, b *, Mikko Heino a, b, c

a Institute of Marine Research, PO Box 1870, Nordnes, NO-5817 Bergen, Norway
b Department of Biology, University of Bergen, PO Box 7803, NO-5020 Bergen, Norway

* Corresponding author, Tel: +47 55 58 42 49, Fax: +47 55 58 44 50

Email address: Devine: ja.devine@ymail.com; Heino: mikko@imr.no

**ABSTRACT**

Changes in size and age at maturation of many exploited fish stocks have been investigated and the influences of environmental factors and exploitation have often been inferred, but not explicitly investigated. Here we determine probabilistic maturation reaction norms (PMRN) for Barents Sea haddock (*Melanogrammus aeglefinus*) using generalized linear models (GLM) and mixed effect models (GLMMs), which account for the correlation among samples within a tow station, and investigate the effects of fishing mortality, environmental factors (NAO, water temperature, and salinity), and potential density dependence or species interaction effects. We found little evidence of a consistent trend in maturation tendencies for Barents Sea haddock for cohorts 1983–2003, ages 4–6 years. Female haddock matured at larger lengths for a given age than males, but overall patterns were similar for both sexes. The GLMM approach gave consistently higher PMRN midpoint estimates than the GLM.
approach, which indicated that PMRNs that do not account for correlations within the data may bias estimates. Environmental factors, rather than exploitation, density dependence, or species interactions, were responsible for the observed changes in size and age at maturation in Barents Sea haddock during the late 1980s through early 2000s. Little evidence of fisheries-induced evolution was found in these fish over the time period investigated. The lack of a significant temporal trend in maturation may be due to several challenges in estimating representative population parameters for this stock, the trait change being in a period of stasis or reversal, or adaptation by the fish to higher exploitation in the past resulting in negligible evolutionary selection during the study period when exploitation has been more moderate.

Keywords: Barents Sea, haddock, life history, maturation, Melanogrammus aeglefinus, probabilistic maturation reaction norms

1. Introduction

Harvesting of wild populations is not random selection. Humans are capable of generating relatively rapid phenotypic changes within harvested populations by targeting individuals of certain sizes, age classes, or behaviours (Hendry et al., 2008; Darimont et al., 2009). Concentrated and high selection pressure on individuals can result in genetic changes in the population if the selected phenotype has a partial genetic basis (Miller, 1957; Ricker, 1981; Nelson and Soulé, 1987; Rijnsdorp, 1993a; Law, 2000). Life history traits that have a partial genetic basis and have been shown to evolve under selective pressures include age and size at maturation (Silliman, 1975; Reznick et al., 1990; Reznick and Ghalambor, 2005), reproductive effort (Rijnsdorp et al., 2005), growth (Conover and Munch, 2002; Conover et
Traits also evolve in response to changing environmental conditions, which include density dependent growth responses resulting from changes in intraspecific or interspecific competition, oceanographic conditions, and changes in prey availability or type (Stokes et al., 1993). Pinpointing whether trait change is a result of phenotypic plasticity or genetic evolution must either depend on molecular genetic analysis or common-garden experiments, which are relatively rare, or weaker approaches, such as regression or reaction norm methods (Dieckmann and Heino, 2007).

Probabilistic maturation reaction norms (PMRNs) have been suggested as a method to disentangle the effects of phenotypic plasticity from genetic effects on maturation (Heino et al., 2002b). By describing the probability of becoming mature as a function of age and size, PMRNs are thought to remove the main effects of varying mortality and juvenile growth rates (Dieckmann and Heino, 2007). However, the PMRN approach has limitations. Reaction norms for age and length at maturation do not account for factors other than growth-related phenotypic plasticity in maturation (Dieckmann and Heino, 2007; Uusi-Heikkilä et al., 2011), nor do they disentangle all effects of growth variability on maturation, neither in theory (Heino and Dieckmann, 2008) nor in practice (Morita and Fukuwaka, 2006; Morita et al., 2009). When possible, other variables should be included in the maturation reaction norm estimation, and some work has investigated higher-dimensional reaction norms, which included the effects of condition and temperature on the maturation process (Baulier et al., 2006; Grift et al., 2007; Mollet et al., 2007; Vainikka et al., 2009a). Yet imperfect disentanglement does not make the PMRN approach invalid. PMRNs remove the plastic effects of varying average juvenile somatic growth rates from the description of the maturation schedule (Dieckmann and Heino, 2007), and this is already an improvement over indices that are sensitive to growth variability, such as maturity ogives. Often unaccounted
effects, such as condition and water temperature, can be seen as contributing “mere” noise to estimation. However, when unaccounted effects show trends, they may introduce spurious trends to PMRNs or mask true trends (Dieckmann and Heino, 2007; Heino and Dieckmann, 2008). As with any other statistical tool, results from PMRN analyses must be interpreted critically.

PMRNs have been used to investigate changes in size and age at maturation for many commercial fish stocks (e.g., Heino and Dieckmann, 2008; Wang et al., 2008; Pardoe et al., 2009; Vainikka et al., 2009a; Vainikka et al., 2009b; van Walraven et al., 2010; Swain, 2011), and the general concession has been that an important factor for the observed changes is exploitation. Furthermore, fishing pressure explained the observed changes in life history traits for the majority of the 37 fish stocks investigated by (Sharpe and Hendry, 2009). Most PMRN studies assessing changes in size and age at maturation have not explicitly evaluated the contribution of environmental factors, although there are a few exceptions (e.g., Heino et al., 2002c; Mollet et al., 2007; Pardoe et al., 2009; van Walraven et al., 2010).

Changes in PMRNs have been or currently are being investigated for two of the main commercial species in the Barents Sea, Northeast Arctic cod (Gadus morhua; Heino et al., 2002c) and capelin (Mallotus villosus; L. Baulier, in prep.). Haddock (Melanogrammus aeglefinus) is the second most important commercial groundfish species in the Barents Sea. Stock dynamics are characterized by large cyclic fluctuations (Olsen et al., 2010) and strong recruitment pulses, thought to be partially linked to water temperatures (ICES, 2008). The exploitation rate of haddock has been variable, with a peak catch of 322,000 t occurring in 1973; catches in the last decade have been in the range of 150,000 – 200,000 t (ICES, 2010). Haddock is currently classified as “harvested sustainably”, although unreported discarding is thought to be an issue (ICES, 2010). The main spawning grounds are located to the west and
south of the Barents Sea, along the Norwegian shelf edge (Solemdal et al., 1989), and mature fish typically migrate from the central and eastern areas of the Barents Sea. Spawning is between March and June, with most fish spawning at the end of April (ICES, 2010). Age at 50% maturity in haddock has fluctuated since 1980, but does not show a consistent trend (ICES, 2010).

Here we investigate whether Barents Sea haddock show evidence of changing age and size at maturation between 1981–2009. The stock has been exploited at a higher rate for several decades prior to investigation (ICES, 2010). We use PMRNs to analyze temporal trends in maturation tendency, where the reaction norm method describes the probability that an immature fish will mature during a given time interval and at a certain size and age (Heino et al., 2002b). Since the approach accounts for most of the effects of phenotypic plasticity in growth, we investigate secondary factors, usually not included in PMRN analyses, that might explain any observed trends; these included the effects of fishing mortality, environmental factors from three areas of the Barents Sea (water temperature and salinity), the North Atlantic Oscillation (NAO) index, and potential density dependence or species interaction effects. We use both generalized linear models and generalized linear mixed models. Generalized linear models are the method typically employed in PMRN analyses because of their ease of use; however, generalized linear mixed models naturally allow accounting for correlations within the data, which can lead to incorrect statistical conclusions if ignored (Smith et al., 2009). We used mixed models to incorporate the knowledge that size, age, and maturation status of fish sampled from the same station in a given year are likely highly correlated and to investigate if ignoring such correlation led to large differences in maturation estimates and trends.

2. Methods
Data were collected by the Institute of Marine Research (Bergen, Norway) during the winter Barents Sea research surveys (late January–early March), in 1981–2009. Immature and mature fish were identified based on the maturation status of their gonads. All fish of maturation stage 1, without evidence of spawning zones on otoliths, were considered immature. Newly matured fish (recruit spawner) and repeat spawners were differentiated on the basis of otolith patterns, or the number of spawning zones apparent on the otolith; this is similar to the procedure used for determining cod (Rollefsen, 1933) and halibut (Hippoglossus hippoglossus; Devold, 1938) recruit spawners. Fish were retained in the analysis if the otolith readability scores were 1 (zones counted and measured with certainty) or 2 (zones may be counted, but not measured).

Analyses were restricted to ages where first-time spawning fish appeared in the data. Data were restricted to cohorts 1983–2003 and ages 4–6 for both sexes on account of too few data for recruit spawners from other cohorts and ages. Approximately 3,000 recruit spawner and 10,300 immature fish measurements were used in the analyses (Table A.1). The survey is thought to disproportionately target the immature fraction of the stock as mature fish are beginning to migrate to spawning areas at this time (Aglen et al., 2005). Therefore the proportions of immature and maturing fish in the survey data may not reflect the true proportions in the population. In this scenario, the length at maturation estimate will be biased high because the distribution is skewed towards immature fish. To account for potential biases in sampling, immature fish were down-weighted by the ratio of the proportion of maturing fish estimated from the ICES maturity ogive for the joint Russian-Norwegian winter survey combined with the Russian autumn survey (ICES, 2009) to that estimated only from the winter survey data. The ogives estimated from the combined autumn and winter survey data
are believed to better represent the proportions of immature and mature fish. The proportion of maturing fish \( m(a) \) was estimated from the maturity ogive \( o(a) \) at age \( a \) and \( a-1 \) as:

\[
m(a) = \frac{o(a) - o(a-1)}{1 - o(a-1)}
\]  

(Barot et al., 2004). If data from the Barents Sea winter survey were representative of the stock, then the ratio would equal one and no weighting would occur.

The amount of area surveyed was standardized from 1981–1993, after which it was expanded and the amount of coverage depended on ice extent, Norwegian access to the Russian EEZ (or vice versa), or Russian involvement in the survey (coverage of Russian waters by Russian vessels; details in Aglen et al., 2005). To account for the changing survey area, the ratio of area covered in a particular year to the maximum area surveyed (all years; ICES, 2009) was used to weight the proportions of maturing fish. Maximum area was from the 1995 survey.

The effects of weighting the data were tested in a sensitivity analysis by 1) ignoring the potential unrepresentative sampling of juvenile relative to mature fish, 2) ignoring the change in areal extent of the survey, and 3) removing both forms of weighting. Weighting terms that accounted for these factors were removed sequentially from the final model and the effects on model fits, residual deviance, the random intercept term, and PMRN midpoints were investigated. The sensitivity analysis was only performed on the generalized linear mixed model.

2.1 Probabilistic maturation reaction norm method

The direct method of estimating PMRNs developed by (Heino et al., 2002b) was adapted for use with generalized linear mixed models (GLMMs). A GLMM with a binomial error distribution and logit link function, a random intercept term, data weights as defined above,
and the Laplace approximation of likelihood were used (Bates and Maechler, 2010). The Laplace approximation allows the use of likelihood-based inference (Bolker et al., 2009), which was necessary for determining model choice. Likelihood ratio tests are adequate for testing fixed effects (model selection) in GLMMs when the ratio of the total sample size to the number of fixed-effect levels being tested is large (Pinheiro and Bates, 2000) and the number of random-effect levels are large (Demidenko, 2004; Bolker et al., 2009). The full mixed model with all possible interactions and a random intercept was the base model. The random intercept term was a unique station + year identifier; stations are pre-defined for the bottom trawl survey in the Barents Sea (see Aglen et al., 2008) for details and station placement. Haddock distribution appears to be based on size and/or age and distributions are relatively consistent over time (Aglen et al., 2000; Aglen et al., 2005; Aglen et al., 2008). The random intercept term implies that the probability of a fish becoming mature at a given station in a given year is correlated with other fish of the same size and age at that station and time, and indicates that increasing the number of samples within a station provides less new information than sample size would normally signify. This correlation has been shown to be positive for maturity ogives (Korsbrekke, 1999) and other parameters (Pennington and Vølstad, 1994). Model fits, standard errors of parameter estimates, and residual plots were also used to assess models.

PMRNs were also estimated with generalized linear models (GLM) for comparison with GLMM models. GLMs treat each fish as an independent sample and any correlations between samples within a station are ignored; this is how data are typically treated when estimating probabilistic maturation reaction norms. Data were under-dispersed for GLM models, thus a quasibinomial error distribution with logit link function was used. The quasi-AIC (QAIC) was estimated as \( \text{Deviance}/\text{Dispersion} + (2 \times (df + 1)) \) where \( df \) is the number of degrees of freedom (Burnham and Anderson, 2002).
For all models, length was included as a continuous variable, while age and cohort were first fit as factors and then as continuous variables. Once a model using only length, age, cohort, and interaction terms was chosen, other explanatory factors were added. Likelihood ratio tests were again used to compare models with additional explanatory terms. Models were also assessed using standard errors of parameter estimates and residual plots.

### 2.2 Randomization tests and confidence intervals

Randomizations were used to investigate if the observed differences for sex, cohort, and age in the maturation reaction norms were real (Barote et al., 2004). For example, the hypothesis that males and females had different PMRNs was tested by creating a new data set, where sex values were randomly assigned to individuals within each cohort and age. Within the new, randomized datasets, the probability of maturing was modelled using GLMs, with a quasibinomial error structure and logit link, and included length as a continuous variable and data weights (as defined above). The randomization procedure was repeated 1000 times and the likelihood ratio F statistic, testing the sex effect, was collected. The effect was considered significant if less than 5% of the randomizations led to a higher value of the test statistic than that computed from the original data. A similar procedure was also used to test whether cohorts or ages had different reaction norms. GLMs, and not GLMMs, were used because while both models gave similar results, GLMs took much less time to run (minutes per model as opposed to days).

Confidence intervals for the length at 50% probability of maturing ($L_{50}$) were also estimated with bootstrapping and randomization approaches. Tows were randomly resampled within a year to recreate a new data set from which reaction norm estimates were derived. This process was repeated to obtain 1000 estimates of the reaction norm, and the distribution of the endpoints was then used to derive the 95% confidence intervals (Manly, 1991).
2.3 Explanatory factors

Environmental data were extracted from databases at the Institute of Marine Research and were chosen to represent habitat of haddock lifestages and their food (Olsen et al., 2010). Oceanographic data from the Fugløya-Bear Island transect (70°30' N, 20°30' E – 74°15' N, 19°10' E, 0–200 m) provided an estimate of temperature and salinity from the influx of Atlantic water and are considered representative of the western Barents Sea (Stiansen et al., 2005). Water temperature from the Vardø-North section (31°13' E, 76°30' N – Vardø) was used to represent the central part of the Barents Sea (Ingvaldsen et al., 2006), while data from the Kola transect (33°00' E, 77°30' N – 72°30' N, 0–200 m) provided an indicator of thermal and salinity conditions for the southern and eastern Barents Sea (Bochkov, 1982). The North Atlantic Oscillation (NAO) winter index was used as a measure of the strength of large-scale atmospheric circulation (www.cgd.ucar.edu/cas/jhurrell/indices.html). Positive anomalies appear to be related to northward shifts of the Gulf Stream, more frequent and intense storms in the vicinity of the Norwegian Sea (Hurrell and Deser, 2009), and greater inflow of warmer water into the Barents Sea (Hurrell and Dickson, 2004). The total estimated biomass of cod, haddock, capelin and ratio of capelin to cod biomass were used to represent potential species interactions. Capelin and cod were included because capelin is the primary fish prey of haddock, while cod are known to include a proportion of small haddock in their diet, especially in years of low capelin biomass (Stiansen et al., 2005; Olsen et al., 2010). The ratio of capelin to cod biomass was used as an indicator of higher predation upon haddock. The direct effect of fishing on reaction norms was tested by including fishing mortality (F, year\(^{-1}\)), which was a proxy for the strength of selection imposed by the fishery. Haddock are regulated by a TAC, minimum bottom trawl mesh size of 135 mm (125 mm in the Russian EEZ), sorting grids since 1997, and a minimum size limit of 44 cm (39 cm in the Russian EEZ; ICES, 2010). Trawls with a mesh range of 135–140 mm have a 50% retention of haddock of
47–53 cm average length and have a selection range, i.e., the length range over which retention increases from 25% to 75%, of 38–66 cm (Halliday et al., 1999; Huse et al., 2000). Because regulations have remained fairly static over the time period of the study, we assumed that size selection should have been similar from year to year.

The “decision” to mature occurs long before the process of spawning begins (Wright, 2007). The maturation process has been suggested to take up to three years for North Sea plaice (Rijnsdorp, 1993b; Grift et al., 2003) and one year for North Sea haddock (Tobin et al., 2010). Because the length of time between “deciding” to mature and spawning is unknown for Barents Sea haddock, the effect of explanatory variables one and two years before first spawning was tested. For fishing mortality, the effect seven years before first spawning was also included (lag = 7); this is the generation time for Barents Sea haddock, approximated using the method of (Froese and Binohlan, 2000) and data from ICES (2010). All explanatory factors were standardized to a mean of zero and a standard deviation of one, thereby creating unit-less indices, rescaled around zero (Fig. 1). Correlations between explanatory factors were tested to ensure highly correlated factors were not included in the same model.

3. Results

3.1 Length-at-age

Average size of immature and recruit-spawning haddock increased with age (Fig. 2). Recruit spawners were, on average, 4 cm larger than immature fish of the same age. Age 4 fish showed the largest size differences between immature and mature fish. All ages generally indicated a stable average size, but showed a sharp increase for cohorts in the mid- to late 1980s, and a decrease in size in the early to mid 1990s. For cohorts from the 1980s, difference in size between immature and recruit spawning fish was less (e.g., age 4, both sexes) or
immature fish were larger than recruit spawners (e.g., age 6 males), which indicated sampling may not have been representative of the population.

3.2 Maturation reaction norm

The sample sizes of recruit spawners were too low to use the full model, which included age, year, cohort, and all the interaction terms (i.e., $age \times cohort + length \times age + length \times cohort + length \times age \times cohort$). GLM and GLMM models that best described the probability of maturing for haddock were the same for fish of each sex:

- **Males**: $\text{logit}(\text{maturation}) \equiv c_0 + c_1 \times length + c_2 \times age + c_3 \times cohort + c_4 \times (age \times cohort)$, and
- **Females**: $\text{logit}(\text{maturation}) \equiv c_0 + c_1 \times length + c_2 \times age + c_3 \times cohort + c_4 \times (length \times age)$,

where length and age were fit as continuous variables, and cohort was included as a factor (i.e., a classifying variable). Including the $age \times cohort$ interaction was needed to detect age-dependent temporal changes in the probability of being mature, while the $length \times age$ interaction in the female model allowed the width of the reaction norm to become age-dependent and detected length-dependent changes in the probability of maturing. Although randomization tests showed that no significant differences between ages existed, adding age to the models improved the fit of the model and was necessary to allow for commonly observed differences between age-classes.

The environmental terms that best described trends in the probability of maturing differed between sexes. The Vardø-North section temperature index two years before the maturation event explained the variation in males (GLMM and GLM results), while water temperature from the Fugløya-Bear Island section one year and salinity from the Kola section two years before the maturation event explained the variation in trends in females (GLMM results; Table 1). Water temperature had a positive effect on maturation whereas salinity had a
negative one (Table 1). The Kola section salinity-only model appeared to be the better GLM model for females (Table 1) and the plotted Lp50 estimates did not differ greatly from the two-environmental term model (except for a few cohorts in age 5 and age 6; Fig. 3), indicating that a more complex model may not be necessary. Exploitation, species interactions, and density dependence did not explain the variation in estimates for either sex. The trend in Lp50 estimates for both GLMM and GLM models were relatively similar, but GLM estimates were generally 1–3 cm smaller than GLMM estimates. In other words, not accounting for the similarity of fish within the same trawl haul meant that Lp50 was underestimated by as much as 8% for males and 4% for females.

Females matured at a larger size for a given age than males. Females were, on average, 7.8 cm larger than males when they attained 50% probability of maturing at age 4, 10.3 cm larger at age 5, and 14.5 cm larger at age 6 when comparing midpoints from the GLMM approach. The size difference between the sexes at maturation was slightly greater when comparing the midpoints from the GLM approach: 10 cm for age 4, 11.9 cm for age 5, and 16.4 cm for age 6.

Estimates of reaction norm midpoints for males and females show a similar pattern; Lp50 estimates were smaller for cohorts before 1985 (Fig. 3). After 1985, PMRN midpoints appeared to oscillate slightly, with lower values in the mid-1990s. A large amount of uncertainty was apparent. The 1987 and 1994 male cohorts were poorly estimated as a result of the low number of recruit spawners sampled (Table A.1). For both sexes at age 6, the models fit poorly for those cohorts where the size of immature fish was equal to or greater than the size of maturing fish. This could indicate that the sampling was not representative of the population’s true size structure. Result from t-tests indicated that the slightly increasing
trend in $L_{p_{50}}$ estimates for both sexes and all ages was not significantly different from zero ($p \geq 0.12$ for all cases; Table A.2).

### 3.3 Sensitivity tests

Removing the down-weighting of immature fish (to account for potential oversampling of immature fish) and the weighting by survey area (to account for changes in the survey) resulted in poorer model fits, less residual deviance explained, and higher variances associated with the random intercept term (Table 1). Models without weights were exceptionally poor fits for both males and females and removal of weighting affected which explanatory terms improved the fit of the model. For example, removing all weighting from the female model resulted in no explanatory terms significantly improving the model.

Accounting for differences in the amount of area surveyed each year appeared to be more important than compensating for the undersampling of maturing fish. However, both weighting terms were needed in the model, as seen in the higher remaining residual deviance and random intercept term variance (Table 1). A higher variance in the random intercept term indicated that the logistic curve estimated for each station had a vastly different intercept than the final weighted model.

Maturation reaction norm midpoints were lower after accounting for potential sampling issues. Ignoring that immature fish were most likely overrepresented in the trawl hauls resulted in midpoints 3–5 cm larger, which was 5–12% larger than midpoints from the model including all forms of weighting. Removing the weighting by survey area increased PMRN midpoints by 1–4 cm (2–8%), while using no model weights resulted in maturation midpoints 8–21% larger for fish of a given age.

### 4. Discussion
We demonstrate that maturation tendencies of Barents Sea haddock vary sexually (i.e., females tend to mature at larger sizes for a given age than males), but show no specific trend over the past three decades. Maturation tendency is variable and appears to reflect, at a temporal lag, patterns in length at age. Increased length at first spawning has been reported in haddock following reduced growth and therefore smaller size at age, usually as a result of high abundance related to periodic strong recruitment pulses (Templeman et al., 1978; Kovtsova, 1993; Korsbrekke, 1999; Korsbrekke, 2003). This may be an indication that, while capturing most of the plastic effects of juvenile growth rate variations from the description of the maturation schedule (Dieckmann and Heino, 2007), not all effects of growth variability on maturation were removed by the PMRNs.

The probabilistic maturation reaction norm approach has suggested evolution in maturation for numerous fish stocks (reviewed in Jørgensen et al., 2007; Heino and Dieckmann, 2008; Sharpe and Hendry, 2009). The majority of these studies have used only age and size to describe the maturation tendency. Here, we have included information on fishing mortality, environmental factors, and inter- and intraspecies interactions within the PMRN model to determine whether this information strengthens the explanatory power of these models.

Important factors for determining maturation in haddock (beyond body length) were western Barents Sea temperature lagged one year (females), central Barents Sea temperature lagged two years (males) and south/eastern Barents Sea salinity lagged two years (females). Water temperature for the two regions was significantly correlated ($r = 0.54$, $p < 0.001$) and trends were similar (Fig. 1a), which may explain why different factors were important for males and females. When we tested the effect of swapping temperature factors for males and females, western Barents Sea water temperature lagged one year was highly significant for males ($p < 0.001$) and central Barents sea temperature lagged two years was significantly related to trends in females ($p = 0.01$), but model fits were better in the final chosen models for each
sex. Temperature has been increasing in the Barents Sea, with some variability, throughout
the period of interest; trends in salinity are similar to trends in temperature. Increased water
temperature can be due to overall higher water temperatures or because there is an increased
inflow of Atlantic water (Loeng et al., 1997), whereas increased salinity may be the result of
greater ice formation and subsequent brine release (Schauer et al., 2002) or increased inflow
of Atlantic water (Stiansen et al., 2005). Atlantic water volume flux has nearly doubled in the
past decade and warmer, more saline water has spread further eastward in the Barents Sea
(ICES, 2008). Increased inflow often results in increased nutrients and subsequent increased
abundance of planktonic organisms (Ottersen and Stenseth, 2001; Stiansen et al., 2005).
Changes in Atlantic inflow may also be reflected in current velocities and effect the transport
of larvae and zooplankton (Loeng et al., 1997). Gjøsæter and Loeng (1987) postulated that
current velocities and resultant transport played a role in large variations in growth of capelin
when water temperatures within the region did not vary annually.

Neither density dependent nor species interaction effects considered here accounted for
significant variability in the PMRN estimates. Environmental conditions have been shown to
override density dependent effects for species near the northern limit of their range (Ottersen
and Loeng, 2000). This is in contrast to what Heino et al. (2002c) found for Northeast Arctic
cod, where maturation in cod was influenced by feeding conditions (the amount of available
prey and high intraspecific competition for that prey), and the findings of Korsbøkke (1999),
who suggested that maturation in haddock was influenced by density dependence. The effect
of fishing mortality on haddock maturation probabilities was explicitly tested within the
reaction norms and was found not to contribute. Using fishing mortality as a proxy for
selection pressure is not as robust as using a metric that describes the size-selectivity of the
fishery. In years of high harvest rates, harvest rates for younger age classes were higher,
indicating age selectivity may exist in the fishery. The relationship between haddock
maturation and selection pressure warrants further investigation and is the next logical step, especially since exploitation was higher in the three decades prior to this study. However, we must reiterate that there is no temporal trend within the maturation reaction norm midpoints in the past twenty years, indicating that there is little evolution occurring.

The motivation for using generalized linear mixed models, instead of the simpler generalized linear models, is that the former allow a natural way to account for within-sample correlations (Smith et al., 2009). Intra-sample correlations lead to effective sample sizes being lower than the nominal ones, and ignoring this can lead to incorrect statistical inference. In our study, however, the results from both approaches did not differ in any essential way: the temporal patterns were similar and the underlying trend was not significantly different from zero. The GLMM approach, however, did give consistently higher PMRN midpoint estimates than the GLM approach. Thus, PMRNs that do not account for correlations within the data or nested design may bias maturation reaction norm midpoints.

Sampling of haddock in the Barents Sea may not be representative of the population. Large concentrations of mature haddock have been observed to often display pelagic behaviour during the winter survey (Aglen et al., 2005), which would result in more immature fish being captured. Mature fish also are often found high in the water column at night (Olsen et al., 2010), which is long in the Barents Sea at the time of the survey. Since the survey operates over 24-hours, it is possible that there are some catchability issues between day and night sampling. Immature fish may be overrepresented in the catches also because mature haddock begin migrating to the spawning grounds at the time of the survey, although it has been noted that these are mainly the older, mature fish (age 7+; ICES, 2010) and not the ages used in this analysis. Sampling of maturing individuals for a given age class was relatively consistent, albeit low, in the analysis. The reaction norm method is considered insensitive to variability in
sampling intensity or gear size selectivity as long as the maturity proportions remain unchanged (Heino et al., 2002a; Wang et al., 2008). Given the behavioural changes of haddock with maturation and the movement of fish to the spawning ground around the time of sampling, it is possible that the maturity proportions were affected. We attempted to correct for this by applying a weighting factor to the numbers of immature fish based on the predicted probability of maturing estimated from maturity ogives.

The changes in length at age of haddock in response to environmental conditions in this study are similar to that of Korsbrekke (1999, 2003), who restricted his analysis to the central region of the Barents Sea, an area consistently covered by the Norwegian Barents Sea groundfish survey in all years. We choose to use all available data and weight by the proportion of maximum survey area covered to downweight years of less coverage. Growth changes presented here were also similar to trends in growth data from the Lofoten survey (Korsbrekke, 2003), which surveys mainly the mature portion of the stock. This provides some evidence that the winter survey sampling may not be inconsistent with stock dynamics.

Most studies investigating long-term trends in maturation reaction norms have revealed changes that are suggestive of fisheries-induced evolution (summarized in Heino and Dieckmann, 2008). Furthermore, other haddock stocks have shown changes suggestive of fisheries-induced evolution (Neuheimer and Taggart, 2010). That Barents Sea haddock is an exception to this pattern was unexpected. The lack of a clear trend might be because consistent maturation changes did not take place during the study period, the trait change is in a period of stasis or reversal, or that changes occurred, but we were unable to detect them. We did not have any a priori reasons to expect no maturation evolution in haddock. A possible explanation is that although no significant trend over time was detected in maturation reaction norms of Barents Sea haddock, strong selection may have been present earlier: fishing
mortality in the 1950s was approximately double that of recent years and has steadily declined (see Fig. 1e). In the thirty years prior to sampling, adaptation to fishing pressure may have occurred; such rapid adaptation to anthropogenic disturbances, over the time span of a few generations of the organism, has been increasingly reported (Kinnison and Hendry, 2001; Hendry et al., 2008; Darimont et al., 2009). Adaptation to past exploitation could have led to a situation in which selection during the study period was absent or too weak to cause significant evolutionary change. Variation and reversals in evolutionary trajectories is common in many contemporary evolution and paleontological studies (Hendry and Kinnison, 1999), and recent experimental work by Conover et al. (2009) suggests that harvested populations possess the ability to recover from the fisheries-induced evolution. The last scenario, that we were unable to detect a change that actually took place, could also apply because several challenges in estimating representative population parameters for this stock existed – primarily as a result of behavioural changes (and resultant gear selectivity issues) and emigration from the study area. Nevertheless, we conclude that the most parsimonious interpretation of our results from the Barents Sea haddock is that maturation in this population has not evolved during the study period.

5. Conclusions

Changes in maturation of Barents Sea haddock were investigated and multiple potential drivers examined included exploitation, water temperature, salinity, the North Atlantic Oscillation, species interactions, and density dependent effects. No significant temporal trends in maturation exist for the 1983–2003 cohorts, and we could not find evidence for exploitation driving changes in maturation. Instead, water temperature and salinity appeared to be the most important variables tested that explained the variation in maturation for both sexes, on top of the effects of age and length. Possible reasons for the lack of trend suggestive
of fisheries-induced evolution may be that we were unable to detect a change given the
challenges in estimating sampling parameters, that adaptation to higher exploitation in the
past resulted in negligible evolutionary selection during the study period when exploitation
has been more moderate, or that change in the trait is in a period of stasis or reversal.

Acknowledgements
We thank the two anonymous referees, whose constructive comments and suggestions greatly
improved the manuscript. This research has been supported by the Specific Targeted Research
Programme FinE (Fisheries-induced Evolution; contract number SSP-2006-044276), funded
through the European Community’s Sixth Framework Programme, a Norwegian Research
Council grant (project 173417/S40), and the Bergen Research Foundation. The article does
not necessarily reflect the views of the European Commission and does not anticipate the
Commission’s future policy in this area.

Appendix A. Supplementary data

References
13/2008, p. 52.


Kinnison, M.T., Hendry, A.P., 2001. The pace of modern life II: from rates of contemporary microevolution to pattern and process. Genetica 112-113, 145-164.


Table 1

Results of GLMM and GLM models and sensitivity tests (GLMM models only) for the direct PMRN method. The p-value is from the likelihood ratio test between the no environmental term model and the model including the environmental term, and coefficient expl. 1 and expl. 2 refer to the coefficients for the environmental terms. Model term abbreviations are: L = length, A = age, C = cohort, A:C is the age×cohort interaction, L:A = length×age interaction term, VNT.2 = Vardø North water temperature lagged 2 years, FBT.1 = Fugløya-Bear Island water temperature lagged 1 year, and KS.2 is salinity from the Kola transect lagged 2 years.

Fig. 1

Standardized plots of explanatory variables used in the analysis: (a) water temperature from the Kola transect, Fugløya-Bear Island (FB) transect, and Vardø-North section (VN); (b) salinity from the Kola and Fugløya-Bear Island (FB) transect; (c) NAO index; (d) biomass of cod, capelin and haddock; and (e) capelin:cod biomass ratio, and haddock fishing mortality averaged for ages 3-7 and 3-11.

Fig. 2

Mean length and standard deviation of male and female haddock by age and cohort. Black circles represent mature fish, open (grey) circles are immature fish.

Fig. 3

Temporal trend in the reaction norm for size and age at maturation midpoints for male and female haddock. Black circles represent GLMM model estimates, grey circles are GLM estimates (open grey circles are the one-term GLM model). Vertical bars are the bootstrapped 95% confidence intervals of the estimates.
<table>
<thead>
<tr>
<th>Model</th>
<th>AIC/ QAIC</th>
<th>Deviance</th>
<th>Random effects: $\sigma^2$</th>
<th>p-value</th>
<th>Coefficient Expl. 1</th>
<th>Coefficient Expl. 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Males</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) $L + A + C + A:C$</td>
<td>GLMM</td>
<td>3398</td>
<td>3310</td>
<td>3.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) $L + A + C + A:C + VNT.2$</td>
<td>GLMM</td>
<td>3390</td>
<td>3300</td>
<td>3.21</td>
<td>&lt; 0.001</td>
<td>1.18</td>
</tr>
<tr>
<td>(1) $L + A + C + A:C$</td>
<td>GLM</td>
<td>7079</td>
<td>3553</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) $L + A + C + A:C + VNT.2$</td>
<td>GLM</td>
<td>6966</td>
<td>3534</td>
<td>–</td>
<td>&lt; 0.001</td>
<td>1.06</td>
</tr>
<tr>
<td>Model (2) without down-weighting</td>
<td>GLMM</td>
<td>3998</td>
<td>3908</td>
<td>9.56</td>
<td>&lt; 0.001</td>
<td>1.36</td>
</tr>
<tr>
<td>Model (2) without survey area weighting</td>
<td>GLMM</td>
<td>4117</td>
<td>4027</td>
<td>12.47</td>
<td>1.47</td>
<td></td>
</tr>
<tr>
<td>Model (2) without any weighting</td>
<td>GLMM</td>
<td>4777</td>
<td>4687</td>
<td>29.20</td>
<td>2.00</td>
<td></td>
</tr>
<tr>
<td><strong>Females</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) $L + A + C + L:A$</td>
<td>GLMM</td>
<td>2782</td>
<td>2732</td>
<td>3.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) $L + A + C + L:A + FBT.1$</td>
<td>GLMM</td>
<td>2773</td>
<td>2721</td>
<td>3.22</td>
<td>&lt; 0.001</td>
<td>0.84</td>
</tr>
<tr>
<td>(5) $L + A + C + L:A + KS.2$</td>
<td>GLMM</td>
<td>2768</td>
<td>2716</td>
<td>3.33</td>
<td>&lt; 0.001</td>
<td>-0.67</td>
</tr>
<tr>
<td>(6) $L + A + C + L:A + FBT.1 + KS.2$</td>
<td>GLMM</td>
<td>2758</td>
<td>2704</td>
<td>3.25</td>
<td>&lt; 0.001</td>
<td>0.86</td>
</tr>
<tr>
<td>(3) $L + A + C + L:A$</td>
<td>GLM</td>
<td>5402</td>
<td>2945</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) $L + A + C + L:A + FBT.1$</td>
<td>GLM</td>
<td>5505</td>
<td>2928</td>
<td>–</td>
<td>&lt; 0.001</td>
<td>0.67</td>
</tr>
<tr>
<td>(5) $L + A + C + L:A + KS.2$</td>
<td>GLM</td>
<td>5266</td>
<td>2926</td>
<td>–</td>
<td>&lt; 0.001</td>
<td>-0.46</td>
</tr>
<tr>
<td>(6) $L + A + C + L:A + FBT.1 + KS.2$</td>
<td>GLM</td>
<td>5333</td>
<td>2909</td>
<td>–</td>
<td>&lt; 0.001</td>
<td>0.68</td>
</tr>
<tr>
<td>Model (6) without down-weighting</td>
<td>GLMM</td>
<td>3238</td>
<td>3184</td>
<td>7.92</td>
<td>0.97</td>
<td>-0.85</td>
</tr>
<tr>
<td>Model (6) without survey area weighting</td>
<td>GLMM</td>
<td>3436</td>
<td>3382</td>
<td>11.34</td>
<td>0.92</td>
<td>-0.92</td>
</tr>
<tr>
<td>Model (6) without any weighting</td>
<td>GLMM</td>
<td>4004</td>
<td>3950</td>
<td>23.46</td>
<td>1.00</td>
<td>-1.16</td>
</tr>
</tbody>
</table>
Figure 3

PMRQ midpoint (cm)

Males, age 4  Males, age 5  Males, age 6

Females, age 4  Females, age 5  Females, age 6

Cohort