Ecological and genetic impact of Atlantic cod larval drift in the Skagerrak

Nils Chr. Stenseth\(^1,2,\)*, Per Erik Jorde\(^1,2\), Kung-Sik Chan\(^3\), Elizabeth Hansen\(^3\), Halvor Knutsen\(^1,2\), Carl André\(^4\), Morten D. Skogen\(^5\) and Kyrre Lekve\(^1\)

\(^1\)Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biology, University of Oslo, PO Box 1066 Blindern, 0316 Oslo, Norway
\(^2\)Institute of Marine Research, Fledlevigen, Nye Fledlevveien 20, 4817 His, Norway
\(^3\)Department of Statistics and Actuarial Science, University of Iowa, Iowa City, IA 52242, USA
\(^4\)Department of Marine Ecology, University of Gothenburg, Tjärnö Marine Biological Laboratory, 452 96 Strömstad, Sweden
\(^5\)Institute of Marine Research, PO Box 1870 Nordnes, 5817 Bergen, Norway

We evaluate the hypothesis that Atlantic cod larvae are passively transported by sea currents from off-shore spawning areas to settle in coastal waters, a hypothesis which has recently gained support from genetic analysis of cod in the North Sea–Skagerrak area. Such larval transport has been suggested to be an important mechanism behind the commonly observed low spatial genetic differentiation in many marine organisms. Here, we apply an ARMAX(2,2) model for juvenile abundance and use long-term monitoring data from the Skagerrak coast, constituting 54 continuous annual series from 1945 to 1997. Analysing the model, we find that the product of the size of the North Sea breeding stock and the strength of the net inflow of North Sea waters had a significant, positive effect on the abundance of coastal juvenile cod. The peak effect occurs during the month of March, just after spawning, when eggs and larvae remain pelagic and sensitive to currents. In contrast, we find no evidence of any direct effect of the North Sea spawning stock alone. Our analyses indicate that 15–20 000 0-group larvae from the North Sea reach each fjord per year, on average. This corresponds to about 1–10% of the total 0-group population in each fjord on average. These findings clearly demonstrate a direct link between larval drift and gene flow in the marine environment.

**Keywords:** genetic differentiation; time series modelling; ecological dynamics; larval drift; Atlantic cod (\textit{Gadus morhua})

1. INTRODUCTION

The marine environment presents ample opportunities for passive transport of pelagic eggs and larvae by ocean currents over large distances. Indirect evidence for the ecological and genetic impact of such larval drift comes from the observation that marine species are often genetically similar over large geographic distances as compared to terrestrial and freshwater organisms (e.g. Waples 1987). However, observations directly linking the existence of a pelagic larval stage to extensive gene flow are scarce, and counter-examples exist (Pogson et al. 1995; Taylor & Hellberg 2003). Here, we use a unique combination of long-term ecological monitoring data, oceanographic modelling, and genetic analyses of Atlantic cod (\textit{Gadus morhua}) in the North Sea–Skagerrak area (figure 1) to evaluate the ecological and genetic impact of larval drift in a marine species.

In a recent study (Knutsen et al. 2004) we demonstrated, by means of DNA microsatellite analyses, that juvenile (young of the year) cod caught along the Norwegian Skagerrak coast in a year with strong inflow of North Sea water (2001) are genetically more similar to mature North Sea cod than they are to adult coastal cod. This situation was reversed in another year (2000) when inflow of water masses from the North Sea was weak (cf. figure 2a,b). Here, we evaluate the hypothesis that the changing genetic profile of coastal juvenile cod may be attributed to (temporally variable) inflow of North Sea cod larvae by extending and applying a previously reported ecological model for cod in the coastal Skagerrak system (Chan et al. 2003a,b).

2. MATERIAL AND METHODS

(a) Data

The nearest major ‘upstream’ spawning ground for cod in the North Sea, from which larvae may potentially reach Skagerrak by passive current drift, is the German Bight. Spawning in this part of the North Sea peaks in early March (Brander 1994) and cod larvae typically settle to the bottom at the end of April and thereafter remain fairly stationary (Julliard et al. 2001; Chan et al. 2003a).

We estimated the annual North Sea spawning biomass from data from the International Bottom Trawl Survey (IBTS: data supplied by ICES) monitoring surveys in the first quarter from 1965 to 2002. The IBTS data consist of catch-per-unit-effort (CPUE) of cod from age group 1 and older,
and these were averaged and weighted by the area of the sampling rectangle in the North Sea region, to yield an estimate of spawning biomass \( B_j \) in year \( t \): \[ B_j = \sum_{i=1}^{n} P_i N_{i,j} W_{i,j}, \] where \( P_i \) is the proportion of mature individuals in age group \( j \) that spawn, \( N_{i,j} \) is the density estimate of the group \( j \) in year \( t \) and \( W_{i,j} \) is the mean body-weight for group \( j \) in year \( t \). The mature proportions and the mean body weights were supplied by ICES (ICES 2001; see figure 3).

A three-dimensional wind- and density-driven ocean model, NORWECOM (the Norwegian Ecological Model System; see Skogen & Søiland 1998), was used to estimate the direction and magnitude of water flux on monthly and daily bases. We modelled the flux of water in the direction from the cod spawning grounds in the North Sea German Bight towards the Skagerrak, across the transect indicated with a dotted line in figure 1, for the period from 1955 to 2001.

The NORWECOM model is forced by wind, air pressure (data obtained from the Norwegian Meteorological Institute at http://met.no/english/index.html), tides and freshwater runoff. Water fluxes (figure 2b) were calculated on a daily basis during the spawning period and the following weeks (March and April), at a spatial resolution of 20 km² for the upper 50 m water column and for salinity less than 35‰; see figure 3b–d. The salinity limit was set to the North Sea average in order to exclude the more saline Atlantic water-masses, whereas the depth was set close to the maximum depth in the study area.

Annual monitoring data on juvenile coastal cod abundance were obtained from 54 locations in the Skagerrak and grouped into 15 fjords based on geographic proximity (figure 1; table S1, electronic supplementary material; see also Danielsen 1969) and covering the period 1945–1997. The overall dynamics of the abundance along the Norwegian Skagerrak coast is depicted in figure 2c for the purpose of illustration, although we used fjord-specific values in the model. These monitoring data represent a unique set of time series; for further information about this sampling program, see for instance Smith et al. (2002). For earlier studies that have used these data for ecological modelling, see Fromentin et al. (1997, 1998, 2000, 2001), Stenseth et al. (1999), Bjørnstad et al. (1999a,b) and Chan et al. (2003a,b). Building upon the ecological model of Chan et al. (2003b), we here extend the application of these monitoring data to testing the population genetic hypothesis suggested by Knutsen et al. (2004).

(b) Statistical modelling

Based on a population dynamics model for the coastal cod along the Norwegian Skagerrak coast, Chan and co-workers (Chan et al. 2003a,b) developed an Auto-Regressive Moving-Average eXogenous variable model, ARMAX(2,2), for the logarithmically transformed 0-group cod abundance series. The model considers a series of coastal locations (or ‘fjords’) to represent demographically (semi-) autonomous populations, and is in its general form given as: \( n_{i,j}^t = f(\text{past abundances and environmental conditions}) + \eta_{i,j}^t \), where \( n_{i,j}^t \) is the log-transformed abundance of 0-group (i.e. young of the year) cod caught in fjord \( j \) in the autumn of year \( t \) (i.e. \( n_{i,j}^t = \log(N_{i,j}^t) \)). The model incorporates within-
Figure 2. (a) Statistical assignment (Knutsen et al. 2004) of cod larvae sampled in years 2000 and 2001: to the North Sea population (black bars); to adult populations at the same location (grey bars); and to other adult Skagerrak populations (white bars). (b) Modelled water flux ($10^6$ m$^3$ s$^{-1}$) from the North Sea into Skagerrak, across the depicted transect with a dotted line in (a), during March and April in years 2000 (open symbols) and 2001 (filled symbols). (c) Population dynamics of juvenile Skagerrak cod (number of cod caught per beach seine haul, averaged over fjords).

Figure 3. The covariate data. (a) The time plot of the logarithmically transformed annual spawning biomass. (b) The time plot of the daily net inflow into Skagerrak in year 1971; days are counted from March 1. (c) The similar plot of net inflow in 1997. (d) The time plots of daily net inflow for all years.
between-cohort interactions, interactions with coexisting species (mainly adult Pollock, Pollachius pollacius), environmental factors such as water temperature and the North Atlantic Oscillation (NAO; Hurrell & Van Loon 1997; Stenseth et al. 2003), experimental larval releases (Smith et al. 2002), as well as the effect of an extensive algal bloom in 1988. The \( \eta_{ij} \) elements are fjord-specific coloured noise terms that are order-2 moving-average processes.

Here, we extend this coastal-cod model by incorporating the effect of the (log-transformed) North Sea breeding stock \( b_t \) (data available since 1971: see ICES 2001) and the strength of the sea current inflow into the Skagerrak (\( C_{2} \) see Skogen & Soiland 1998) during the time from spawning and several weeks afterwards (March–April; Brander 1994). This extension was done by introducing larval-drift effects into the model so that it now takes the form

\[
\tilde{n}_{j,t} = f(\text{past abundances and environmental conditions}) + (\zeta_j + \phi_j C_{t})b_t + \eta_{ij,t},
\]

where \( C_t \) is the net flow of North Sea water into the Skagerrak during March (see below) in year \( t \). Because the spawning biomass data are available only since 1971, we used a two-stage approach for estimating the model so as to utilize all information. We adopted the approach of eliminating the effects due to intra- and inter-specific factors by replacing \( n_{j,t}^{0} \) by the residuals \( (\tilde{n}_{j,t}^{0}) \) from the earlier model, which did not include the term \((\zeta_j + \phi_j C_{t})b_t\). We thus analysed the residuals of the earlier model, introducing a direct effect \((\zeta_j \text{ due to migration})\) and an indirect effect \((\phi_j \text{ due to pelagic eggs or larvae being carried with the sea current})\) on coastal juvenile cod yielding the following generic, fjord-specific, model:

\[
\tilde{n}_{j,t}^{0} = k_j + \text{DME} + \text{ILdE} + \epsilon_{ij,t},
\]

where \( \text{DME} = \zeta_j b_t \) is the direct effect of mature North Atlantic cod migrating/swimming into the Norwegian Skagerrak coast and \( \text{ILdE} = \phi_j C_{t} b_t \) (an expression which will be further developed below). Furthermore, \( k_j \) is the intercept (equal to 0 if the covariates are of zero mean and we had been able to analyse the whole set of residuals from the original model, and not just those from 1971) and \( \epsilon_{ij,t} \) are temporally uncorrelated noise terms that may have contemporaneous (between-fjord) correlation, thereby accounting for spatial correlation non-parametrically. Thus modelled, any direct effect of the North Sea cod would indicate that mature North Sea cod migrate into the Norwegian coastal localities ("fjords") and spawn there (a possible, though unsubstantiated proposition), whereas an indirect effect would indicate that sea currents \((C)\) flowing into Skagerrak transport larvae produced by the North Sea spawning stock \((B = \exp(b))\)

In order to focus on the parameters of interest to this study \((\zeta_j \text{ and } \phi_j)\) we first eliminated the intra- and inter-specific effects \((k_j)\) by replacing \( n_{j,t}^{0} \) with the residuals \( (\tilde{n}_{j,t}^{0}) \) of the earlier model of Chan et al. (2003a,b). Two sets of analyses of the residual model were undertaken: in the first set we assumed that the parameters \( \zeta_j \text{ and } \phi_j \) differ between fjords, but are fixed over time; in the second set \( \zeta_j \text{ and } \phi_j \) were assumed to be identical for all fjords (an assumption emerging as a conclusion from the first set of analyses), but \( \phi_j \) is allowed to vary during the course of the season (thus, we add a subscript, \( d \), for the day since March 1, prior to the North Sea spawning season (Brander 1994); i.e. \( \phi_d \)).

(i) Model 1: spatially variable, time-invariant effect parameters

In this model variant we allow the parameters to vary between fjords but keep them fixed during the course of the season. For each fjord \( d \) we estimated the parameters \( \kappa_d, \zeta_d \text{ and } \phi_d \) in the model using Proc Model of the statistical software SAS via the method of iterated seemingly unrelated regressions (Hamilton 1994) which allows for a general contemporaneous covariance matrix for the stochastic error terms; this non-parametrically models the spatial correlation structure of the process.

The parameters of the model—assuming them to be time-invariant—were estimated from annual \((i)\) data on juvenile cod abundance \([n_{j,t}^{0} = \log(N_{j,t})]\), North Sea spawning stock size \([b_t = \log(B_t)]\), and inflow \((C_t)\) of North Sea water to the Skagerrak coast. In this set of analyses we assumed ILdE to be given by \( \phi_j C_{l} b_t \).

(ii) Model 1b: can the effect be seen as an effect of food (plankton) being transported into Skagerrak?

The effect we see of ocean currents on coastal cod could, in principle, be caused by inflow of food particles (zooplankton) leading to increased survival of locally produced coastal cod, rather than by direct influx of North Sea larvae. We thus evaluate the following alternative model: \( n_{j,t}^{0} = f(\text{past abundances and environmental conditions}) + \gamma C_t + \eta_{ij,t} \), where \( \gamma \) represents the (common) direct effect of the sea current onto the local-fjord cod populations along the Norwegian Skagerrak coast. This direct effect could represent zooplankton produced in the North Sea being transported by the sea current into the Skagerrak area. The effect during the latter part of March is indeed significant (see table S2, online supplementary material), which is also the case for the \( \phi_d \) for the same period. However, it is not significant for April, which it should be if this represents a real effect of plankton. Thus, the strength of the current over the second half of March in this particular model is most reasonably seen as a proxy for the product of the log-transformed (true) spawning biomass and sea-current. Presumably, there are enough eggs and larvae in the North Sea in any year to yield an impact on the coast when currents are favourable. Also, there is likely to be much noise in the estimated spawning biomass density variable, this being supported by the fact that the \((\zeta_j + \phi_j C_{l} b_t)\)-model indeed provides the statistically better fit. This overall conclusion is supported by Beaugrand et al. (2003), who documented high concentrations of zooplankton for a substantially longer period than the latter part of March. Altogether we conclude that the \( \gamma C_t\)-model is inappropriate for our system. Hence, in the rest of this paper we use the model \( n_{j,t}^{0} = f(\text{past abundances and environmental conditions}) + (\zeta_j + \phi_j C_{l} b_t + \eta_{ij,t}) \) or the variety of it presented below in the paper.

(iii) Model 2: spatially uniform and temporally variable effect parameters

In this model variant we assume common direct and indirect effect parameters, but allow indirect effects to vary during the course of the season. The indirect effect, ILdE, is assumed to be given as \( \sum_{t=1}^{d} \psi_{j} \phi_{j,t} b_t \), where \( t \) denotes the year and \( d \) the day since March 1, and the summation is taken over the entire time window during which larvae could potentially reach the Skagerrak coast (until the end of April). The parameter \( \psi_{j} \) is proportional to the probability that a (random) mature cod spawns on day \( d \). Estimating the \( \psi_{j} \) without any constraints on them may not be feasible or result in very erratic estimates because of relatively high ratio of parameter per data point. However, the \( \psi_{j} \) are expected to be a ‘smooth’ function of \( d \), \( d=1,2,\ldots,61 \), in that the second differences \( \psi_{j,2} = 2\psi_{j,1} + \psi_{j,-1} \) should be close to 0, where \( \psi_{j,2} \) are specified as 0 for \( d \) less
than 1 or larger than 61. All the parameters, including $\phi_s$, are estimated using a penalized weighted least squares method that balances the goodness of fit and the smoothness of the $\phi_s$, where the smoothness is determined by the method of generalized cross validation. Details are provided in Hansen et al. (2004). The smoothness constraints lead to the estimated $\phi_d$ curves in figure 4a,b having effective degrees of freedom equal to 1.9 and 1.5, respectively.

In our preliminary analysis, we found that the appropriate time window for which to include the water-flux from the North Sea to the Skagerrak in the model is March, especially the latter half of March, and possibly early April (cf. table 1; see electronic supplementary material for detailed description).

3. RESULTS

The first set of analyses reported in table 2, assuming time-invariant parameters (model variant 1), suggests that except for fjord no. 8 (the Risør skerries) there is no significant difference between fjords with respect to the $\zeta_j$ and $\phi_j$ parameters (representing the direct and indirect effect, respectively, of the North Sea cod stocks on the coastal Skagerrak populations). We thus calculated common fixed values (model variant 2) for the eight fjords for which the parameters appeared consistent (i.e. all fjords except the Risør skerries; see table 1 and electronic supplementary material for details).

For the assessment reported in table 1, we evaluated three measures of ‘inflow’: ‘inflow’=inflow (inflow being positive by convention), ‘net inflow’=inflow+outflow (outflow being negative by convention) and ‘truncated net inflow’=max(inflow+outflow,0). The use of truncated net inflow is motivated by the fact that only positive inflow may lead to contribution from the North Sea to the Skagerrak cod systems. For each of the inflow measures, we consider seven time-windows or periods, and calculate the mean during the given period. Table 1 lists the $p$-values for the direct North Sea cod spawning biomass effect, $\zeta$, and the indirect effect, $\phi$, of spawning North Sea cod and currents (i.e. the ILdE ($=\phi_j C_{j\beta}$) term in our model) on the abundance of 0-group cod along the Norwegian Skagerrak coast. As some estimates of $\phi$ are negative, the $p$-values for $\phi$ are those corresponding to one-sided tests of the null hypothesis of $\phi$ being zero versus the alternative hypothesis of $\phi$ being positive. Eight fjords (1, 4, 5, 6, 7, 10, 11 and 14; see legend to figure 1) are used and common parameters are assumed for them.

All three measures of inflow gave similar inferences. While the use of net inflow leads to more significant effects, it also gives rise to an anomalous significant direct effect $\zeta$ in period IV (cf. table 1) which, however, is deemed insignificant using the other two measures of the sea current. Henceforth, we use the net inflow in March as a measure of the effective magnitude of current flowing into the Skagerrak.

Using the water inflow from the North Sea to Skagerrak during the latter part of March, the inclusion of all model parameters was found appropriate for 9 of the 15 fjords in the sense that $R^2$ values for these fjords are positive. Specifically, three of the estimated direct ($\zeta_j$) effects are significant (at the nominal 5% level): however, two of them are negative. The indirect parameter ($\phi_j$) was significant and positive for three fjords, consistent with the hypothesis of passive transport of North Sea larvae into the fjords.

In order to increase statistical power we also performed a combined analysis over eight fjords by assuming that the parameters $\zeta_j$ and $\phi_j$ have common values ($\zeta$ and $\phi$) across fjords. We used eight fjords in this analysis, rather than nine, because as indicated above, location 8 (Risør skerries) deviates from the other eight fjords, and probably does not share a common structure with the other fjords.

Excluding Risør yields a significant indirect effect ($\phi=0.11$: s.e. = 0.05; $p=0.02$), suggesting that North Sea cod influence coastal populations indirectly. The significant $\phi$ value most probably represents an inflow of larvae from the North Sea to the Norwegian Skagerrak fjords, in accordance with the findings of the genetic analysis (Knutsen et al. 2004). Further, we find no significant direct effect of the adult North Sea cod biomass on the Skagerrak juvenile populations ($\zeta = -0.11$: s.e. = 0.07; $p=0.14$), a result agreeing with previously reported tagging studies of...
Table 1. Assessing the ‘optimal’ time-window for calculating the inflow of North Sea waters (i.e. the time period providing the strongest signal on the 0-group cod at the Norwegian Skagerrak coast). (The table lists one-sided $p$-values for the common direct North Sea cod spawning mass effect, $\zeta$, and the common indirect effect (North Sea cod spawning biomass) $\times$ (current), $\phi$, influencing the abundance of 0-group cod along the Norwegian Skagerrak coast. Eight fjords (1, 4, 5, 6, 7, 10, 11 and 14; see legends to figure 1) are used and common parameters are assumed for them. Here, ‘net inflow’ is net inflow + outflow (outflow being negative by convention) and ‘truncated net inflow’ is max(inflow + outflow, 0). The use of truncated net inflow is motivated by the fact that only positive inflow may lead to contribution from the North Sea to the Skagerrak cod systems.)

<table>
<thead>
<tr>
<th>period</th>
<th>$p$-values for inflow</th>
<th>$p$-values for net inflow</th>
<th>$p$-values for truncated net inflow</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>direct effect of North Sea cod</td>
<td>indirect effect of North Sea cod</td>
<td>direct effect of North Sea cod</td>
</tr>
<tr>
<td>I (March 1–15)</td>
<td>0.152</td>
<td>0.071</td>
<td>0.134</td>
</tr>
<tr>
<td>II (March 16–31)</td>
<td>0.120</td>
<td>0.039</td>
<td>0.176</td>
</tr>
<tr>
<td>III (April 1–15)</td>
<td>0.188</td>
<td>0.318</td>
<td>0.132</td>
</tr>
<tr>
<td>IV (April 16–30)</td>
<td>0.347</td>
<td>0.996</td>
<td>0.025</td>
</tr>
<tr>
<td>V (March 1–31)</td>
<td>0.124</td>
<td>0.034</td>
<td>0.139</td>
</tr>
<tr>
<td>VI (March 16 to April 15)</td>
<td>0.095</td>
<td>0.087</td>
<td>0.117</td>
</tr>
<tr>
<td>VII (April 1–30)</td>
<td>0.455</td>
<td>0.873</td>
<td>0.181</td>
</tr>
</tbody>
</table>

Table 2. Likelihood ratio tests of the hypotheses of common parameters $\zeta$ and $\phi$, across nine fjords. ((a) Nine fjords (1, Torvefjord; 4, Bufjord; 5, Fledevigen; 6, Sandnesfjord; 7, Søndeledfjord; 8, Risør skerries; 10, Kilsfjord; 11, Soppekilen and 14, Vestfjord). (b) Eight fjords (same as in a but with no. 8, Risør skerries, excluded). The hypothesis of common parameters among the remaining fjords is clearly supported when excluding fjord 8.)

<table>
<thead>
<tr>
<th>test value ($\chi^2$)</th>
<th>d.f.</th>
<th>Pr &gt; $\chi^2$</th>
<th>comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) nine fjords included</td>
<td>39.09</td>
<td>8</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>17.81</td>
<td>8</td>
<td>&lt; 0.0227</td>
<td>common $\phi$</td>
</tr>
<tr>
<td>(b) eight fjords included (the same as in part b except for the Risør skerries)</td>
<td>10.79</td>
<td>7</td>
<td>0.148</td>
</tr>
<tr>
<td>6.61</td>
<td>7</td>
<td>0.471</td>
<td>common $\phi$</td>
</tr>
</tbody>
</table>

4. DISCUSSION AND CONCLUSION

Our ecological result clearly supports the hypothesis put forth by Knutsen et al. (2004) of cod larvae being carried with ocean currents into the coastal areas (‘fjords’) in the Skagerrak. Taken together, our new results, together with the former genetic analyses, strongly suggest that larval drift is the primary cause for the observed high genetic similarity among cod populations in the North Sea–Skagerrak area, rather than adult dispersal or other factors.

The scenario of dispersal of cod larva from a common donor population into multiple recipient populations can be approximated by a standard infinite ‘island’ model. This allows for some rough quantification of the number of North Sea immigrants required to explain the low amount of genetic differentiation observed among cod populations in the Skagerrak–North Sea system (average $F_{ST}$ = 0.0023 (s.e. = 0.0006); Knutsen et al. 2003). The estimate $N_{m} = 1/(1/F_{ST} – 1)/4 = 108$ (95% confidence interval from 71 to 227) represents the average number of immigrants per generation from the North Sea stock that enters the coastal populations and survive until adulthood.

While often problematic (Whitlock & McCauley 1999), such an estimate based on $F_{ST}$ appears reasonable in the present situation. First, the observed ‘star’ phylogeny among Skagerrak cod populations (Knutsen et al. 2003) is consistent with gene flow following the theoretical island model, as expected if genes are introduced into the coastal populations from a common, presumably large, source.


