Fish assemblages inside and outside marine protected areas off northern Iceland: protection effects or environmental confounds?

Andrzej Jaworski¹,*, Jon Solmundssonb, Stefan Aki Ragnarssonb

¹ Marine Scotland – Science, Marine Laboratory, PO Box 101, 375 Victoria Road, Aberdeen, AB11 9DB, UK
² Marine Research Institute, PO Box 1390, Skulagata 4, 121 Reykjavik, Iceland

Abstract

The density and mean size of demersal fish were analysed inside and outside three protected areas off the northern coasts of Iceland. One-way ANCOVA and mixed effects models, with depth, bottom temperature and tow duration as covariates, were used to determine differences in the two metrics between areas. In general, these differences were not statistically significant, except for one protected area in the northeast. In this area, small size classes of haddock (Melanogrammus aeglefinus) and long rough dab (Hippoglossoides platessoides) were found to be considerably more abundant and large size classes of cod (Gadus morhua) and haddock less abundant compared to the reference (fished) area. The mean sizes of haddock and long rough dab were smaller within this protected area. While these differences could not be confidently attributed to the closure, this area may be considered particularly suitable for the protection of juvenile fish. The large effect of covariates (mainly that of depth and temperature) on observed densities and mean size reasserts the necessity of their inclusion in models assessing the impact of area closures on fish communities. The study illustrates that a snapshot approach is unlikely to provide unambiguous evidence of the effectiveness of area closures. It can, however, provide useful information on various relationships between environmental gradients and fish distribution inside and outside a closure.

1. Introduction

It is widely recognised that marine protected areas (MPAs) are an important regulatory measure to improve the management of human impacts on the marine environment (Horwood et al., 1998; Polunin, 2002; Gell and Roberts, 2003; Kaiser, 2005; Stefansson and Rosenberg, 2005; Jones, 2007; Jennings, 2009). Evaluation of their effectiveness, i.e. determining whether they actually meet their objectives, is a fundamental part of their implementation. It allows prompt modification of management and monitoring strategies (Gerber et al., 2005). More generally, such evaluation provides information to support management decisions within the framework of adaptive management (Agardy, 2000; Gerber et al., 2005).

Strict evaluation of the effectiveness of MPAs is difficult because in most cases, the objectives and criteria of success are poorly defined (Polunin, 2002; Willis et al., 2003). Moreover, availability of high-quality empirical data, such as time series, which would demonstrate the effectiveness of MPAs, is often limited. Empirical evidence for recovery in MPAs frequently suffers from a lack of rigour in the design of field surveys (Mosquera et al., 2000; Willis et al., 2003). Willis et al. (2003) identified the main problems associated with insufficient replication and a lack of control sites in evaluation studies. In the past, many sampling programmes for evaluation were initiated long after the establishment of MPAs (Pelletier et al., 2008), making it more difficult to ascertain whether observed differences between a given protected area and reference (fished) areas were an effect of the area closure or whether they existed before the closure. On the other hand, it may be difficult to distinguish changes in fish populations caused by natural processes in the environment from those induced by the reduction of fishing effort. As effects of area closures may be mitigated by a whole range of factors and processes (Polunin, 2002), detecting such effects is not guaranteed even with a valid survey design. Willis et al. (2003) stressed that little evidence exists to substantiate responses of fish populations to area closures. They noted, however, that this does not imply that MPAs fail in their objectives.

In Iceland, area closures are an important measure to protect fish stocks in addition to allocating quotas on species and mesh size regulations (Jaworski et al., 2006). Most closures (mainly off northwest, north and east Iceland) aim at protecting juvenile fish to increase the long-term yield to the fishery. Some areas are subject to temporary or seasonal closures, while others are closed permanently. In addition, some major spawning areas for the main commercial stocks are closed during spawning time. Despite the large number of closed areas in Iceland, their effect on fish populations has...
only recently been studied. Jaworski et al. (2006) investigated the impact of two area closures off the northeast and southeast coasts of Iceland on the demersal fish community. In that study, data from groundfish surveys carried out annually over a 20-year period were used to examine differences in density, mean size and fish species diversity between closed and fished areas, and between periods before and after the closure. Both closed areas had an effect on density and mean size of some commercially exploited fish species.

The present study explores, at relatively small spatial scales, differences between demersal fish assemblages in protected and fished sites off the north coast of Iceland. More specifically, its objective is to establish whether there were significant differences in fish density and mean size between protected and adjacent fished areas and whether any differences in these two metrics could be attributed to the protection status of the areas. The suitability and usefulness of such comparative analyses is discussed in the context of development, implementation and evaluation of MPAs.

2. Material and methods

2.1. Study areas

The data for this study were collected during two surveys conducted by the personnel of the Marine Research Institute aboard RV “Bjarni Saemundsson”. The first survey was carried out during 11 days in August 2004 off the northwest coast of Iceland (Fig. 1a), in the protected area “Northeast of Horn” (P1) and an adjacent reference area (R1). The second survey was carried out during 10 days in July 2005 off the northeast coast (Fig. 1b) in two protected areas, “Off the Northeast Coast” (P2a) and Langesgrunn (P2b), and also in an adjacent reference area (R2).

All three protected areas were established as part of a network of MPAs in the region (Fig. 1) to protect juvenile fish, mainly cod (Gadus morhua). They were closed for trawling and fishing with longlines in 1993 (Fig. 1c and d). However, area P2a has been closed for trawling since the early 1970s, and areas P1 and P2b since 1992. The first two areas also varied in shape before the closure in 1993, especially area P2a. Some experimental trawling (for purposes other than the present study) took place in the eastern part of area P1 between August and November of 2001 (Fig. 1c; note also that the status of the large area west of area P1 was changed from permanent to a seasonal closure in 1998).

2.2. Sampling design

Sampling stations were selected to cover large parts of the protected areas. The original criteria for the location of sampling stations in the reference areas were twofold. First, sampling in the reference areas aimed at keeping environmental factors (bottom and surface temperature, and depth) similar to those in the protected areas. Second, sampling aimed at covering grounds with considerable fishing effort (Fig. 1c and d). Due to strong environmental gradients present in the study areas, it was difficult to select more remote reference areas that would satisfy the first cri-
The choice of sampling stations outside the protected areas was a compromise to match both criteria as closely as possible. This resulted in the sampling stations outside the protected areas being located at a relatively short distance from the borders of the closures. The summary of data collected and the ranges of environmental factors in the study areas is given in Table 1.

A standardised groundfish survey trawl was used in both surveys. The gear had the following specifications: 31 m long headline, 18.3 m long bobbin footrope (rockhopper) that weighed about 2.25 t, 64 m long bridles, 1.95 t otter boards, and 41 mm mesh size. The trawl was towed over the bottom at a speed of 2.9 knots. Tow duration in the two surveys varied widely, particularly during the 2004 survey (Table 1). In the analysis, catch data from both surveys were expressed as numbers of fish per nautical mile. This index was assumed to be a relative measure of fish density.

The data collected for individual hauls included time, geographical position, bottom depth, bottom and surface temperatures, fish species and their numbers at length. In a few sampling stations in the northwest, only cod were measured (Table 1 and Fig. 1a). The analysis was conducted separately for each of the two surveys, individually for the most abundant species (with average densities >1 fish per meter per km) and also for the aggregate category of commercial species. The latter refers to marketable species in Iceland (for their close-to-complete list see Jaworski et al., 2006).

Only sparse data on seabed type, e.g. multibeam data, were available from other surveys for the northeast study area and some data from benthic sampling (photographs and grab data) for the two study areas. However, for most of the sampling stations in this study these data were not available. Consequently, no information on bottom characteristics or benthic habitats was incorporated into the analysis.

2.3. Data analysis

A one-way ANCOVA model, with bottom depth, bottom temperature and tow duration as covariates, was used to determine the significance of any differences in fish density by size class between areas within each study area. In most cases, 10-cm size classes were used. Narrower (5 cm) size classes were chosen for two smaller-sized species, Norway pout (Trisopterus esmarkii) and long rough dab (Hippoglossoides platessoides). Density data were log-transformed [ln(n + 1)] before analysis to achieve variance homogeneity (Gunderson, 1993).

A linear mixed effects model was used to analyse the significance of any differences in mean size between areas. Here, “area” was treated as a fixed factor (as in the ANCOVA model for densities), while “tow” (a grouping factor for fish) was a random component. The covariates (depth, temperature and tow duration) were level-2 predictors (they varied only at the tow level). Only a random intercept was assumed in the mixed effects model (Zuur et al., 2007).

The main factor in the ANCOVA and mixed effects models was “area”, which had two levels (P1 and R1 in the northwest) or three levels (P2a, P2b and R2 in the northeast). Bottom depth and temperature were incorporated into the models either as quadratic or linear terms. The second order polynomial was only employed when the resulting parabola was concave downward, which suggested a unimodal species response along the gradient. Tow duration was included in the models to account for its possible effects on catch rate, particularly on catch rate by size class (and thus also on mean size). The effect of tow duration was assumed to be linear. No interaction terms were considered in the ANCOVA or mixed effects models. Multiple comparison procedures (Tukey–Kramer method for unequal sample sizes) were used as post hoc tests to determine differences between the three areas in the northeast. Any differences found in the statistical analyses with $p<0.05$ were considered as significant.

The percentage of variation explained by the ANCOVA and mixed effects models was calculated for each species. In the ANCOVA models, this variation was represented by the traditional $R^2$ statistic. In the mixed effects models, the $R^2$ statistic proposed by Vonesh and Chinchilli (1997) was used, which measures the proportionate reduction in residual variation explained by fixed effects (Edwards et al., 2008). The relative importance of each explanatory variable in the ANCOVA model was assessed using variance partitioning (Legendre and Legendre, 1998; Zuur et al., 2007). The amount of variation purely related to each explanatory variable (thus not shared with the other variables in the model) was calculated. No such quantification was attempted here with the mixed effects models for the sake of their complexity. All analyses in the present study were conducted using the statistical package R (R Development Core Team, 2009).

3. Results

The most abundant species off the northwest coast were (in decreasing order) Norway pout, haddock (Melanogrammus aeglefinus), cod, long rough dab, redfish (Sebastes marinus), saithe (Pollachius virens), starry ray (Amblyraja radiata) and Atlantic wolffish (Anarhichas lupus). They made up 99% of the survey catch (in numbers) in this study area. Cod, long rough dab, haddock, redfish and Atlantic wolffish were the most abundant species off the northeast coast and constituted 96% of the survey catch there.

Off the northwest coast, the apparent differences in density between areas (Fig. 2a and solid lines in Fig. 3a) were, in most cases, not significant in models with covariates (dotted lines in Fig. 3a). For example, Norway pout and haddock were, ignoring other effects, more abundant in area P1 than in area R1, but the two species also tended to be more abundant at higher temperatures (Fig. 4), which were more prevalent in the protected area (mean 5.4°C) than in the reference area (mean 3.7°C, Table 1). In contrast, cod tended to be more abundant at the low temperature range (Fig. 4) and
Fig. 2. Mean log numbers (observed values by size class) of fish per nm in the protected and reference areas off the northwest (a) and northeast (b) coasts. Size distributions in protected areas are shown in black (areas P1 and P2a) or dark grey (area P2b), and those in reference areas (R1 and R2) in light grey. The vertical dotted lines show significant differences between areas in the ANCOVA model.

their observed density was higher outside the protected area. The width of the confidence intervals for the model-estimated differences varied among species and size classes (Fig. 3a). There was greater uncertainty (about the conclusion of no difference) for more abundant species, for which a high spatial variability was observed (e.g. Norway pout, haddock, cod and long rough dab), and fairly low uncertainty for less abundant species (e.g. saithe, starry ray and Atlantic wolffish). In the latter case, the confidence interval fell roughly within the range of a twofold difference (after antilog transformation) in density between areas.

In the northeast case, where fish were sampled within a narrower range of depth and temperature, marked differences in density per size class were found for some species between area P2a and the two remaining areas, R2 and P2b, whereas the two latter, in general, did not differ significantly (Figs. 2b and 3b). The observed differences between areas were similar to those estimated from the model with covariates (Fig. 3b). Small haddock and long rough dab (≤30 cm in length) were significantly more abundant in area P2a (roughly six times), and large cod and haddock (>40 cm in length) significantly less abundant (roughly nine times) compared to areas R2 and P2b. This shift in response between 30 and 40 cm was also evident for the aggregate category of commercial species (in which cod, long rough dab and haddock made up 83% of the catch in numbers in the northeast).

The observed differences in mean size were in most cases not statistically significant (Fig. 5). In the northwest, only the mean size of Atlantic wolffish was significantly smaller inside the protected area (by 12 cm). The differences in haddock density between areas in the northeast (with small haddock being more abundant and large haddock less abundant inside area P2a, Fig. 2b) resulted in the mean size of haddock being significantly smaller inside the protected area P2a compared to areas P2b and R2 (by 16 cm; Fig. 5, lower panel). Similarly, long rough dab were smaller in area P2a compared to areas P2b and R2 (by 5 cm). Also, commercial fish species taken together were significantly smaller inside area P2a (by 12 cm).

For individual species, the ANCOVA model explained on average from 11 (Atlantic wolffish in the northeast) to 45% (haddock in the northwest) of the variability in density (Fig. 6). The variable “area” explained a considerable portion (for some species, the largest por-
Fig. 3. Differences, observed (solid line) and estimated from the ANCOVA model (dotted line) in mean log numbers of fish per nm between inside and outside the closed area: (a) off the northwest coast (between areas P1 and R1) and (b) off the northeast coast (between areas P2a and R2). The values above the horizontal axis indicate higher fish densities inside and those below indicate higher fish densities outside. The grey bands show 95% confidence intervals for the model-estimated differences. Note the different scales on the y-axes.

Fig. 4. Relationship between fish density (observed log numbers of all sizes) and two environmental variables, bottom depth and temperature, for three selected species off the northwest coast (black = protected area P1, light grey = reference area R1).
Fig. 5. Mean size of fish (with SE) in the protected (P1, P2a and P2b) and reference (R1 and R2) areas off the northwest (NW) and northeast (NE) coasts. The dotted lines show significant differences between areas in the mixed effects model.

Fig. 6. Proportion of variance explained by different explanatory variables in ANCOVA models (only their pure effects) for fish density off the northwest (NW) and northeast (NE) coasts. Note that the proportion of variance explained is shown as the mean of corresponding proportions calculated for individual size classes within each species/category.

4. Discussion

We used a snapshot approach to explore whether there were differences in the demersal fish assemblage between protected and adjacent fished areas at small spatial scales. Although marked differences in fish density and mean size were observed between protected and reference areas, they were in many cases due to varying depth or temperature, rather than a true effect of the area status. In one area, where the differences between inside and outside were statistically and biologically significant, we were unable to unambiguously attribute them to the effect of closure.

The large effect of covariates found in the present study reasserts that they should be considered in evaluations of the impact of area closures, conducted based on fish metrics. Their inclusion in the analysis makes detection of any real effects of a closure more likely, and failure to include them in the analysis runs the risk of misinterpreting the effects of studied closures. This is particularly the case when influential explanatory variables vary widely. As the present study shows, environmental variables may in some cases be important even though their observed ranges are relatively narrow (e.g. depth in the study area in the northeast). In this study, depth and temperature appeared on the whole to be influential covariates (see also Ottersen et al., 1998; Welch et al., 1998; Falco et al., 2007). These two environmental variables appeared weakly intercorrelated in this study, thus indicating no serious problem with collinearity.

Apart from environmental factors, tow duration, if varying widely from tow to tow, may also be considered an important factor (Godø et al., 1990; Somerton et al., 2002; Battaglia et al., 2006). Although detected in some cases in the present study, this effect proved difficult to describe quantitatively. It seems that estimates of fish density and mean size (or any other similar metric) are more likely to be unbiased with uniform tow duration. The
dependence of gear efficiency upon depth (with swept area and vertical opening varying with depth) may be a potential issue when considering density estimates. We consider any bias potentially introduced by this dependence in the present study as negligible, with the relatively narrow depth ranges here, compared to those in other studies, where this effect was significant (e.g. Godø and Engås, 1989).

One important limitation of this study was the lack of suitable data on some other important environmental factors such as bottom type or structure of benthic habitats (Pelletier et al., 2008), which could be crucial for small fish (Benaka, 1999). Put another way, the observed differences (especially those in the study area in the northeast) could be due to factors other than the closure itself, which we could not quantify and account for in our analysis.

The model-estimated differences in fish density between area P2a and the adjacent fished area were statistically significant and large in magnitude. Although it is likely that protection was driving these differences, more evidence is needed before this conclusion can be confidently drawn. The uncertainty here is linked to the fact that area P2a also differed significantly from the other protected area in the northeast, P2b, for which no effect was found compared with the reference area. However, it has to be remembered that area P2a was protected for as long as three decades, and P2b for just over a decade prior to the survey. As the meta-analysis conducted by Claudet et al. (2008) shows, positive effects of marine reserves on commercial fish species are linked to the time elapsed since the establishment of protection. Further, the size structure of the fish assemblage inside area P2a, with considerably more abundant small size classes and less abundant large size classes (by many orders of magnitude), indicates a well-established pattern, supporting the view that protection might have been the causal mechanism for these differences. The lower mean size found in area P2a seems to contradict a common objective of MPAs: to increase fish sizes in catches. Nevertheless, it has to be borne in mind that...
Fig. 8. Effect of temperature (a) and tow duration (b) on mean size of fish in the areas off the northwest coast. Circles denote observed mean sizes (for individual tows), adjusted to mean depth and mean tow duration (a) or mean depth and mean temperature (b). Lines show predicted trends from mixed effects models (with the remaining covariates kept constant at their mean values). Other symbols and notations as in Fig. 7.

The specific objective of this closure was to protect juvenile fish. Our results suggest that this objective was achieved. The long duration of the closure in area P2a may have considerably increased the habitat complexity there, which in turn promotes survival of juvenile fish (Auster and Malatesta, 1995). The pattern observed in area P2a suggests that juvenile fish there may be less exposed to predation pressure from large fish (see also Jaworski et al., 2006). Finally, the fishing effort by otter trawlers outside area P2a was low for several years preceding the survey, which may indicate that it was a less attractive fishing ground to trawlers compared to more offshore fishing grounds (Fig. 1d). Taken together, area P2a may be considered particularly suitable for the protection of juvenile fish, thus potentially reducing their bycatch (Rijnsdorp and van Beek, 1991).

The lack of detectable effects in the two other protected areas (P1 and P2b), despite high fishing effort in adjacent fished areas, might have been due to the relatively short duration of these closures, compared to area P2a, or to fish movement between protected and fished areas (Polunin, 2002). In the first case, a relatively long period of recovery can be expected for long-lived species such as most species in this study (10–20 years according to Polunin, 1997; see also Jennings, 2001). Furthermore, there are strong spatial environmental gradients (in temperature and salinity) over the shelf north of Iceland, particularly in the northwest, where relatively warm southerly water masses meet colder water from the north (Jonsson and Valdimarsson, 2005). In this area, the large spatial variability in environmental conditions affects many fish populations (Astthorsson et al., 2007). The detection of any effects in such variable conditions within time frames comparable to recovery times in other closed areas may be difficult. It should also be noted that if benthic habitats are particularly important for fish, the recovery times for fish may occur on longer time scales. In the presence of considerable fish movement, re-considering the location and extent of the closures may be an option to meet the goals that they were designed for. Nevertheless, such migrations do not imply the ineffectiveness of MPAs (assumed to be properly designed) as a management tool (Stefansson and Rosenberg, 2005).

The lack of a significant difference between areas in the northwest could also be due to the inadequate sample size. Nevertheless, for a number of fish species there, especially those less abundant in the survey catch, the confidence limits for the difference in density between areas were within a twofold difference range, which is considered as the range of biological indifference in terms of abundance (Edgar and Barrett, 1997). Thus, a bigger sample size would not lead to the conclusion of a biologically important difference. For other species (those more abundant and with more variable densities), a bigger sample size would obviously have given more power to the statistical test.

Adequate reference areas are difficult to find, since environmental conditions in protected and reference areas must be matched closely, but yet these areas should preferably be unaffected by each other (Gell and Roberts, 2003; Halpern et al., 2004). Due to the aforementioned high spatial variability in environmental con-
ditions in north Icelandic waters and the distribution of trawling effort (with very low effort in some areas, Fig. 1c and d), we were not left with many options in choosing reference areas. The reference areas in this study were located near the borders of the protected areas, and fish movement between them must be considered likely. Temperature and food availability are factors that could swiftly alter fish distribution within the study areas, and such changes may have affected our results. Similarly, a potential spillover of fish from protected areas to the adjacent fishing grounds would make it more difficult to observe positive effects of the closures (Halpern et al., 2004). Spillover effects could potentially be studied using more reference areas located further away from the protected areas and being comparable in terms of environmental conditions (Halpern et al., 2004), but this, as mentioned previously, proved difficult to achieve in our study.

The analyses in the present study were based on data collected, in each study area, within short time intervals, and thus with no possibility of studying the dynamics of the assemblage, which is important when evaluating the effectiveness of area closures. One reason for using a snapshot approach was that the number of hauls taken in each study area was considerably higher than the number of hauls taken locally during annual groundfish surveys in a single year. Also, sampling stations in such surveys are not located to address the questions posed in this study. Nevertheless, it is likely that a Before-After-Control-Impact (BACI) assessment design (as in Piet and Rijnsdorp, 1998; Frank et al., 2000; Ferraris et al., 2005; Jaworski et al., 2006) using time series of survey data would provide more conclusive evidence of the effectiveness of the studied closed areas (Polunin, 2002; Pelletier et al., 2008). Sweeting and Polunin (2005) concluded that without a BACI design, effects of MPAs can only be inferred rather than attributed. Nevertheless, both types of analysis, a snapshot approach comparing inside vs. outside (as in the present study) and a BACI approach, require adequate reference areas, or preferably a set of randomly selected reference areas (Underwood, 1992).

The present study illustrates that a snapshot approach is unlikely to provide unambiguous evidence of the effectiveness of area closures, especially when environmental factors and habitat characteristics differ between the areas being compared and are not accounted for properly in the analysis. However, in our opinion, such analyses are meaningful (provided that they are sufficiently robust) as their results may provide a representative picture of fish communities inside and outside closures. If repeated, such studies could become part of a monitoring programme. Owing to the small scale in our study, we were able to accurately investigate various relationships between environmental gradients and fish densities/size distributions. This information is also of larger scales. A comprehensive knowledge of fish distribution in relation to important environmental factors (such as depth, temperature and habitat type), but also in relation to the spatial distribution of fishing effort is crucial—first, in the phase of planning and implementation of fishery closures, and subsequently, in their monitoring and evaluation to support management decisions.

Acknowledgements

We thank the crew and researchers aboard RV "Bjarni Saemundsson" for their assistance during sampling. We also thank Peter Wright and three anonymous referees for their comments on earlier versions of this paper.

References


Benaka, L.R. (Ed.), 1999. Fish Habitat: Exclusive Fish Habitat and Rehabilitation. American Fisheries Society, Bethesda, Maryland, USA.


Somerton, D.A., Otto, R.S., Syrlaja, S.E., 2002. Can changes in tow duration on bottom trawl surveys lead to changes in CPUE and mean size? Fish. Res. 55, 63–70.


