Scottish Government
Riaghaltas na h-Alba
gov.scot

Using commercial and survey data to infer real-time fish distribution in the North Sea at high resolution

Fishing Industry Science Alliance (FISA Project 01/15)

Scottish Marine and Freshwater Science Vol 8 No 10

C T Marshall, R Wiff and T Cornulier

marine scotland

Fishing Industry Science Alliance (FISA)

Using Commercial and Survey Data to Infer Real-Time Fish Distribution in the North Sea at High Resolution

Scottish Marine and Freshwater Science Vol 8 No 10

C. Tara Marshall, Rodrigo Wiff and Thomas Cornulier

Published by Marine Scotland Science
ISSN: 2043-7722
DOI: 10.7489/1973-1

Marine Scotland is the directorate of the Scottish Government responsible for the integrated management of Scotland's seas. Marine Scotland Science (formerly Fisheries Research Services) provides expert scientific and technical advice on marine and fisheries issues. Scottish Marine and Freshwater Science is a series of reports that publishes the results of marine and freshwater scientific work that has been carried out for Marine Scotland under external commission. These reports are not subject to formal external peer-review.

This report presents the results of marine and freshwater scientific work carried out for Marine Scotland under external commission.

Using Commercial and Survey Data to Infer Real-Time Fish Distribution in the North Sea at High Resolution

FISA 01/15

C. Tara Marshall, Rodrigo Wiff and Thomas Cornulier
University of Aberdeen

Table of Contents

1 Executive Summary

The EU Landings Obligation has focussed attention on the urgent need to develop effective strategies for reducing the catch of unwanted species or sizes of fish. Real-time reporting is the term used for the rapid collation, analysis and dissemination of bycatch data so as to enable skippers to improve the match between catch composition and available quota. Informed by experience with bycatch reduction in US fisheries, this report (FISA 01/15) considers how real-time reporting could be used in Scotland and outlines a workplan for developing this capacity.

In Scotland, there are several sources of data that are useful for real-time reporting. E-logbook information is currently available in near real-time to individual producer organisations and used for internal reporting purposes. Ongoing improvements to software will soon make these real-time data more accessible to industry. Bycatch information is also collected by the observer programmes coordinated by the Scottish Fishermen's Federation and Marine Scotland Science. Fisheries-independent data are available from surveys conducted twice a year.

Using juvenile cod in the North Sea as an example of unwanted bycatch, observer data were merged with survey data in a scientifically robust statistical framework to develop maps of juvenile cod distribution by month which have the potential for mapping densities of unwanted bycatch with high spatial and temporal resolution. A Bayesian modelling approach was used such that the model could be continually updated in time as new information became available.

The use of real-time reporting in Alaskan and Pacific Northwest demersal fisheries to meet regulatory limits on bycatch of salmon was reviewed. Data about the location and magnitude of salmon bycatch are shared in real-time across fishing vessels belonging to the same fishing cooperative. High bycatch triggers e-mail alerts which are sent to skippers who then use the information for tactical decision making. The information is also used by cooperative managers to establish area closures, termed rolling hotspots, and monitor effectiveness of these closures.

Consultations with Scottish industry revealed general agreement about the utility of real-time reporting for bycatch reduction but reservations about the likelihood of getting skippers to share information. It was also apparent that some skippers are already sharing information across a small network of peers via social media. Developing incentives for sharing information about bycatch within a trusted network of skippers, for example those belonging to the same producer organisation, is critically important.

Real-time reporting utilises existing data resources and available computer and information technology to enhance spatial selectivity. Recommendations are made regarding the further development of statistical models and real-time reporting systems. The need for institutional and attitudinal change is highlighted.

2 List of Acronyms

Acronym	Definition
AFA	American Fisheries Act
BSAI	Bering Sea and Aleutian Islands
CCCS	Cod Conservation Credit Scheme
DATRAS	ICES Database of Trawl Surveys
EBS	Eastern Bering Sea
EM	Electronic Monitoring
EMFF	European Marine and Fisheries Fund
FIS	Fisheries Innovation Scotland
FISA	Fishing Industry Science Alliance
FMP	Fisheries Management Plan
GAM	Generalized Additive Models
GIS	Geographic Information System
GMRF	Gaussian Markov Random Field
GRF	Gaussian Random Field
HSCC	High Seas Catchers' Cooperative
IBTS	ICES International Bottom Trawl Survey
ICA	Intercooperative Agreement
ICES	International Council for the Exploration of the Sea
ICT	Information and Communication Technology
INLA	Integrated Nested Laplace Approximation
IPA	Incentive Plan Agreement
ITQ	Individual Transferable Quota
LO	Landings Obligation
MCMC	Markov Chain Monte Carlo
MCRS	Minimum Conservation Reference Size
MSS	Marine Scotland Science
NMFS	National Marine Fisheries Service
NPFMC	North Pacific Fisheries Management Council
PCC	Pollock Conservation Cooperative
PO	Producer Organisation
PSC	Prohibited Species Catch
RHS	Rolling Hotspot
RTC	Real Time Closures
SFF	Scottish Fishermen's Federation
SIDI	Scottish Industry Discards Initiative
SSAP	Salmon Saving Area Plan
VMS	Vessel Monitoring Scheme

3 Background

Unwanted bycatch and discards have been serious concerns globally, posing a threat to the sustainability of fisheries through economic, biological and ecological losses. There are two general approaches to avoiding unwanted catches: enhanced spatial selectivity (for example 'moving on' rules), and enhanced gear selectivity. Spatial selectivity, or avoidance, determines where and when vessels fish, whereas gear selectivity determines how vessels catch fish. Spatial avoidance measures have been used throughout the world to mitigate for bycatch and discarding (Little et al. 2014). Spatial avoidance is particularly well-developed in the west coast of North America where it has been employed successfully in demersal fisheries of the Bering Sea (Madsen and Haflinger 2015) and the Pacific Northwest (Sylvia et al. 2014). In several eastern Bering Sea (EBS) fisheries, bycatch data are shared in real-time such that "hotspots", defined by their high bycatch rates, can be identified quickly and the information disseminated rapidly to the fishing fleet. In this way, individual vessels have the option of selectively avoiding these areas. This implementation of spatial selectivity is a cost-effective way to facilitate bycatch reduction but requires high levels of fleet participation (Gauvin et al. 1996; O'Keefe et al. 2010; Bethoney et al. 2013). Furthermore, bycatch species should be unevenly distributed in space to be able to identify hotspots, a condition that is met in the EBS by species as halibut and red king crab but not by Chinook salmon.

In the context of Scottish fisheries, spatial avoidance of choke species through real-time reporting of catch rates is a potential discard mitigation method. There are several relevant sources of geo-referenced data that are useful for tracking the catch of choke species. In Scotland, e-logbook information is held by Marine Scotland. It is available to individual producer organisations (PO) for reporting purposes and has a high degree of temporal resolution about catch rates of choke species. Spatial resolution of these data is limited to the International Council for the Exploration of the Sea (ICES) statistical rectangle (gridded by latitudinal rows with intervals of 30', and longitudinal columns with intervals of 1°). More highly resolved data on the location of these hauls is available from Vessel Monitoring System (VMS) data. When combined, these two fisheries-dependent data streams could give levels of spatio-temporal resolution that are comparable to those being used for hotspot mapping in the EBS fisheries. Relevant information includes the data collected by the observer programme coordinated by the Scottish Fishermen's Federation (SFF), and the corresponding Marine Scotland Science (MSS) observer programmes on-board the demersal fleet. The MSS and SFF observer programmes cover Scottish demersal and *Nephrops* vessels, and the majority of vessels grant access to observers where possible. For both programmes the number of observer trips undertaken in 2014 and 2015 is given in Table 1.1. The MSS observer data also has high definition data on vessel position, activity and, in some cases, footage from the closed circuit TV (CCTV) programme. In addition, fisheries-independent data are available from the International Bottom Trawl Surveys (IBTS) conducted twice a year and collated by ICES.

Table 1.1: The number of trips in the MSS and SFF observer programmes. The bracketed values indicate the percentage of all fishing trips covered by the observer programmes.

Year	Number of MSS Observer Trips	Number of SFF Observer Trips
2014	71 (0.27%)	144 (0.56%)
2015	80 (0.33%)	135 (0.56%)

The high degree of spatio-temporal resolution inherent in these different data streams has the potential to improve spatial selectivity of North Sea demersal stocks, including implementation of avoidance measures for unwanted species and size classes. However, the difficulty with combining different data types (fisheries-dependent and -independent) lies in accounting for intrinsic differences in sampling methodology, including units of measurement, gear type, availability, catchability, and sampling strategy (i.e., distribution of hauls in space and time).

This Fishing Industry Science Alliance (FISA) project was conceived originally as a spatio-temporal fish distribution model that utilised novel statistical approaches to combining fisheries-dependent and –independent data. A Bayesian approach was taken such that the model could be continually updated in time as new information became available. In the course of developing a computationally intensive statistical model it became obvious that a practical means of delivering real-time data to the model was required. This led directly to discussions with the individuals doing the data compilation, analysis and dissemination in the US (Sea State Inc., Seattle, WA). From the real-world perspective of discard mitigation, the model only becomes useful when effective real-time reporting exists. This is not presently the case in Scottish fisheries.

In the European context the Scottish fishing industry were early adopters of spatial selectivity in the form of real-time closures as part of the Cod Conservation Credit Scheme (CCCS; Holmes *et al.* 2011). This experience with spatial selectivity is relevant to real-time reporting and spatial selectivity. The experience of the CCCS showed that a component of the demersal fleet was both adaptive and innovative. The forthcoming Brexit could provide an opportunity to consider custom-built solutions including real-time reporting and enhanced spatial selectivity.

3.1 Project Aims

The first aim of this project (FISA 01/15) was to merge commercial and scientific data sources in a scientifically robust analytical framework. This modelling exercise constitutes a prototype model which is a first-step towards the medium-term goal of creating highly-resolved spatio-temporal maps having predictive capability for unwanted catch including juveniles and choke species. If a protocol for sharing information about catch of unwanted species or size classes in real-time became available, then it would be possible to disseminate updated maps of real-time distribution such that fishers can access them remotely. Therefore, the second aim of FISA 01/15 was to review the experience with real-time reporting in the Alaskan and Pacific Northwest groundfish fisheries where hotspot mapping tools have become essential components of bycatch reduction and fisheries management.

To meet this second aim Aberdeen staff (Tara Marshall, Thomas Cornulier) went to Seattle in December 2015 to discuss the real-time reporting of bycatch with Karl Haflinger of Sea State Inc. Sea State Inc is the “third party” agent contracted by the industry to monitor and report on the salmon bycatch as well as establish area closures (referred to as rolling hotspots) and monitor their effectiveness. The meeting participants also included Steve Martell (then of the International Pacific Halibut Commission, currently of Sea State Inc) and Edward Richardson (At-Sea Processors Association, morning only). Discussions focussed primarily on the experience with EBS Pollock fishery. Key similarities and differences between this fishery and the Scottish demersal fishery (see Section 5.6).

3.2 Project Deliverables

1. Static maps of desired variables, e.g., expected catch under Minimum Conservation Reference Size (MCRS) for species identified as priorities by the industry.

2. A model, coded in the statistical programming language R, for a functional spatio-temporal distribution model for a key commercial species over a range of years having observer data from both industry and MSS (see Appendix 1).

3. A final report including:
- 3a) a description of the spatio-temporal distribution model as well as an evaluation of its performance characteristics (see Section 2);
- 3b) a summary of international experience implementing spatio-temporal distribution models in management of a Pacific groundfish fishery (see Section 3);
- 3c) a summary of consultation with industry regarding practical aspects of collecting relevant data and disseminating spatio-temporal distribution models (see Section 4).

The final report also has a summary of recommendations (Section 5) based on work undertaken to meet all three of the deliverables.

4 Bayesian Spatio-Temporal Modelling of Juvenile Cod

4.1 Introduction

Spatio-temporal models are used to explicitly represent species distribution in space and time. Models considering space and environmental covariates have been applied to investigate the outbreaks of disease in a localised area over time (Blangiardo & Cameletti 2015), model patchy distributions of demersal species (San Martin et al., 2014) and determine hotspots of small sharks (Jaureguizar et al. 2016). In fisheries science, a popular methodology for modelling spatial distribution of species density is geostatistics, which often uses observations from fisheries-independent surveys. Geostatistics is a powerful technique for inference and prediction because takes into account the structural spatial correlation in the variance of the observed variable (Rivoirard et al. 2000). However, the use of geostatistics only provides estimates of spatial predictions at a single point in time (snapshot) without incorporating the underlying temporal dynamics of the population. The time dimension can be incorporated by ordering spatial predictions at different points in time. The complexity of simultaneously modelling both space and time is related to two main issues. Firstly, space need to be explicitly connected in time using statistical models that include both spatial and temporal autocorrelation. Secondly, fisheries data usually involves intensive sampling over large areas either pseudo-randomly (research surveys) or non-randomly (fishing vessels). Depending on the degree of resolution, modelling space and time at once becomes extremely computational intensive (Blangiardo & Cameletti 2015).

Spatio-temporal models have been useful in predicting distributions of species that are clustered in space and time and vary significantly from year-to-year, as is the case of unwanted species in the catch composition (Cosandey-Godin et al. 2015). Unwanted species or size classes are often referred to as bycatch. Bycatch can either be retained and landed but is often discarded at sea for legal or economic reasons (Catchpole et al. 2005). Fisheries bycatch are a serious conservation concern globally and regulation and requirement to reduce bycatch are included in the management regulation of a growing number of nations (Little et al. 2015). Landing obligations are currently being implemented in regional seas in the European Union. By 2019, they will cover all species for which there is a quota and discarding of such quota species will be restricted (European Commission 2015, Scottish Government 2012). In contrast to past fishing practices, fish below the MCRS will have to be landed but cannot be sold for human consumption. Therefore, there is an interest for fishermen and managers to develop effective bycatch mitigation strategies to decrease the probability of catching unwanted species or size classes. In addition to changes in fishing gear regulations, time/area closures have been employed to reduce discarding in the EU (Bailey et al. 2010, Little et al. 2015). However, closing areas does not always result in bycatch reduction, because of the poor match between the closed area and the spatio-temporal distribution of the species that is the object of protection (Bailey et al. 2010).

Bycatch is usually clustered in space and time and also varies significantly from year-to-year (Cosandey-Godin et al. 2015, Jaureguizar et al. 2016). When non-target species vary at predictable space-time intervals (e.g., during spawning migrations), time-area fishery closures can greatly facilitate fisheries' compliance with bycatch regulations (Stram and Ianelli 2015, Madsen and Haflinger 2015). Establishing closed areas requires highly resolved descriptions of the spatio-temporal dynamics of bycatch such as can be collected by scientific observers onboard commercial operations (Wiff et al. 2016). Fishery-dependent data potentially provide an intensive spatio-temporal sampling of fish distribution, compared to fisheries-independent surveys that are restricted in temporal coverage. However, fisheries-dependent data are associated with challenging statistical features including a high proportion of zero observations and non-random sampling which generates both spatial and temporal correlation (Ciannelli et al. 2008, Cosandey-Godin et al. 2015). In addition, fishery-dependent data are influenced by uncontrolled factors including fishing gear configuration, quota availability, trip duration, target intention, environmental variables and market conditions (Pelletier & Ferraris, 2000; Wiff et al. 2008). Data obtained from fisheries-independent research surveys have several advantages over fisheries-dependent data, because sampling location is determined independently of local abundance. This results in unbiased population estimates (Roa & Niklitschek 2007). Consequently, fisheries-independent data is more appropriate for developing a mechanistic model for describing spatio-temporal dynamics of bycatch (e.g., Kristensen et al. 2015).

These stark differences in statistical properties of data originating from commercial fisheries and research surveys pose a challenge to developing density estimators through the merging of these two data sources. Integrated stock assessment combines the two data types via different likelihood functions. Catch rates from commercial data are usually influenced by several factors, and the influence of these factors are isolated using a process known as effort standardisation. Following the same data merging principles underpinning effort standardisation (CPUE), bycatch observations from survey data can be treated as a particular fleet in which fishing operations occurs in certain areas in particular seasons with homogeneous fishing gear. Thus, commercial and survey data can be combined in a similar manner to combining data from different fleet types having different gears and fishing operations and therefore variable effort.

From a statistical viewpoint, generalized additive models (GAMs) are often used to model bycatch (e.g. Ortiz & Arocha 2004) because they incorporate different explanatory variables with flexibility about assumption about the error distribution. GAMs allow a straightforward implementation of the spatio-temporal component by including bi-variate smooth functions for two geographical coordinates (longitude and latitude) and time being treated as fixed factor (e.g., San Martin et al. 2013, Jaureguizar et al. 2016). One of the disadvantages of using GAMs to model the spatial component is that the geometry of the domain is difficult to include in the analysis (Kristensen et al. 2015). GAMs usually include latitude and longitude coordinates, incorporated as fixed effects, and thus do not explicitly include the spatial correlation structure (Cosandey-Godin et al. 2015). In contrast, in classical geostatistics the geometry of the spatial dimension is intrinsically incorporated using a variogram function, accounting for spatial correlation among neighbouring data points.

An alternative of the models presented above are the mixed effect models, in which space and covariates can be combined in a sum of fixed and stochastic effects. The most widely used is known as Gaussian Random Field (GRF). This random field is a stochastic process that essentially represents all spatially explicit processes that may have an effect on bycatch attributes (Cosandey-Godin et al. 2015). GRF has advantages over GAMs or classic geostatistics because GRF conceives the spatial correlation as a stochastic process and allows representation of observations and unobserved (latent) variables, permitting us to account for all uncertainty in with the entire bycatch phenomena. In addition, GRF is based on structured correlation among neighbouring data points which allow to model fine scale processes in comparison with the relatively large spatial resolution achieved using splines via GAM.

GRF for spatio-temporal analyses is usually implemented using hierarchical Bayesian models in which the posterior distribution is approximated using Markov Chain Monte Carlo (MCMC) which is computationally intensive (Gilks et al. 1996). However, a new statistical approach is now ready available, namely Integrated Nested Laplace Approximation (INLA) via the R-INLA package (hppt://www.r-inla.org). INLA is a powerful methodology which approximates via inference the posterior distribution in Bayesian analysis, thus avoiding the computational demand, convergence and mixing problems associated with MCMC analysis (Rue et al. 2009)

Advances in modelling fine spatio-temporal correlation using GRF and covariates provides a powerful framework for modelling bycatch. The aim of this paper is to model the spatio-temporal dynamic of juvenile cod (*Gadus morhua*, where juveniles are defined as <35 cm) in the North Sea using Bayesian hierarchical models with R-INLA. This study develops a modelling framework addressing two highly current topics in fisheries management in the era of the landings obligations: the development of models aimed for real-time spatial management and how non-standard sources of fishery data can be coupled with scientific survey data for generating inferences in a spatio-temporal context. The proposed model provides inference for the spatio-temporal dynamics of juvenile cod combining both survey and commercial fishing data while accounting for different gear types. Results have the potential to improve real-time spatial management by mapping fine spatio-temporal scales dynamics of undersized cod likely to contribute to bycatch problems.

4.2 Materials and Methods

4.2.1 Survey Data

The North Sea International Bottom Trawl-survey (NS-IBTS) is an international demersal trawl-survey conducted in quarter 1 and quarter 3 of each year. NS-IBTS are multispecies surveys with standardised data sampling and processing design and provide data for estimation of relative abundances for groundfish species in an area within 51 N to 62 N latitude and 4 W and 9 E longitude and depths shallower than 300 m (ICES 2012). Haul duration was standardised to 0.5 hours. In this article we used data from 2011 to 2015 downloaded from ICES (DATRAS: http://www.ices.dk/marine-data/data-portals/Pages/DATRAS.aspx). Survey hauls with null observations of juvenile cod, were treated as true zeros in the model. The database used comprised 1110 hauls between 2011 and 2015. Each haul has information regarding geographical position, date, and number caught per length class measured to the nearest centimetre.

4.2.2 Commercial Fishing Data

The commercial fishing data used here were collected as part of discard monitoring programs conducted by Marine Scotland Science (MSS) since 1978 and Scottish Fishermen's Federation (SFF), who started sampling commercial catches more recently. Both programs used a common sampling protocol conducted by on-board scientific observers. Under the European Union Data Collection Framework (EC Regulation 199/2008), EU member states are compelled to collect such data, to supply information on discards for stock assessment purposes. Discard data was acquired from staff at the Marine Lab of Marine Scotland Science (Scottish Fisheries Management Database (FMD) operated by Marine Scotland (www.scotland.gov.uk)). The selection of vessels to place observers on follows a randomization process, but skippers can decline having observers on-board, resulting in a quasi-random sampling of the fleet. Samples are taken from the hauls, and the number of fish is raised to haul-level by multiplying the ratio of the discards in the sample with the total weight of the catch in that haul. In addition to numbers caught by length-class, species and haul, environmental data and information of the gear type is recorded. Only the discarded portion of the catch is georeferenced to haul level (and the total weight of the catch, however the species composition of the catch-weight is unknown). It is therefore more challenging to model the entire length-span of a species using this dataset, as it will require additional assumptions and estimations to determine where the fish above legal landing-size were caught, as landings data (i.e. only the landed component of the catch) is only georeferenced by trip or once per day, and not on a haul level. The discard observer Program covers the waters west of Scotland (ICES division VIa) as well. However, to match the coverage of survey- and observer data spatially, only observer data for the North Sea were used. In total, this added up to 2759 hauls between 2011 and 2015, from multiple trawling fishing fleets targeting different species using nine different fishing gears.

4.2.3 Statistical Model

For clarity, the following description of the statistical model is divided into three sub-sections: (1) general model, (2) exploratory analysis and (3) the structured spatio-temporal model. Subsection 2.2.3.1 describes the general statistical modelling of the count data using Markov Gaussian Random Field (GMRF) with Bayesian inference. Sub-section 2.2.3.2 describes the analysis underpinning the selection of variables to be included in the general model. We also described computation issues in modelling all variables directly in the general model and how we pre-computed some variables in order to make computation tractable. The last sub-section (2.2.3.3) describes the structured spatio-temporal analysis using an intrinsic autoregressive model (also known as "Besag" model) with a regressive temporal structure.

4.2.3.1 General Model

Let us assume that our region of interest is divided into *n* non-overlapping areas. Further, let y_i denote the number of juvenile cod (< 35 cm) in the region *i* distributed such that:

$$y_i \sim Poisson[\lambda_i]$$

where λ_i is a parameter defining the expected rate of cod per exposure time (*i.e.,* trawling time) in the region *i* such as $\lambda_i := E(y_i|\mathbf{x})$. We assumed the existence of a n-dimensional Gaussian field $\mathbf{x} = \{x_i : i \in \mathbb{N}\}$ to be point-wise observed through n_d conditional independent data $\mathbf{y}$. The covariance matrix, the Gaussian field $\mathbf{x}$ and the likelihood model for $y_i|\mathbf{x}$ are all controlled by some unknown hyperparameter θ. In this case, and according to (Martino and Rue 2010), the posterior reads:

$$\pi(\mathbf{x}, \theta|\mathbf{y}) \propto \pi(\theta)\, \pi(\mathbf{x}|\theta)\, \prod \pi(y_i|x_i, \theta). \qquad (1)$$

A latent Gaussian model can be expressed as an additive regression model (Martino and Rue 2010). Our interest here is modelling the number of juvenile cod per trawling time t_i, such $E(y_i|\mathbf{x})/t_i$. The structured, additive predictor $\log(E(y_i|\mathbf{x}))$ accounts for the effects of various covariates in an additive manner.

$$log[E(y_i|\mathbf{x})] = \log(t_i) + \beta_0 + \sum_{\gamma=1}^{n_f} f^{(\gamma)}(c_{\gamma i}) + \sum_{\alpha=1}^{n_\beta} \beta_\alpha z_{\alpha i} + \varepsilon_i \qquad (2)$$

Here, the $\{\beta_\alpha\}$'s represent the linear effect of covariates **z.** The $\{f^{(\gamma)}(\cdot)\}$s are the unknown functions of the covariates **c.** β_0 is an intercept, $\log(t_i)$ represents an offset of the exposure time (trawling time) and ε_i is an error term. Hence, the vector of latent effect is $\mathbf{x} = \{\{log[E(y_i|\mathbf{x})]\}, \beta_0, \{\beta_\alpha\}, \dots\}$. The distribution of observations y_i will depend on the latent effect **x** and a number of hyperparameters θ because **x** is a GMRF (Bivand et al 2015).

$\{f^{(\gamma)}(\cdot)\}$ can take many forms. Here, we are interested in modelling the spatio-temporal dependence and group-specific, random effect for vessels. The spatial domain included the North Sea from 48° to 61° N. This observation window was discretised into *n*=2850 grid cells corresponding to an ICES grid cell 1/16th of latitudinal degree. This model was implemented using R-INLA package (http://www.r-inla.org).

4.2.3.2 Exploratory Analysis

We assumed that spatio-temporal variation in juvenile cod is broadly affected by a combination of average and seasonal predictable structures at large spatial scales and unpredictable short-term structures at a range of spatial and temporal scales (summarized in Table 2.1). "Seasonal effects" were defined as those spatial patterns which vary across months but in a repeatable way across years. Two important components of predictable seasonal effects were: 1) the systematic monthly variation in juvenile cod relative abundance explained by depth, here termed the "depth effect"; and 2) a pure spatial effect that was repeatable across years but not explained by depth, here termed "seasonal spatio-temporal" effect. Unpredictable variation of fish abundance around the predicted seasonal average spatial distribution was partitioned into three components: 1) a fixed year effect accounting for year-to-year stock fluctuations; 2) a fixed month effect shared across years; and 3) a pure spatio-temporal random field for each month by year combination (or "layer"), termed the "spatio-temporal" effect. This third component is conceptually the most important effect because it captures the small-scale structure in the month-specific spatial distribution of fish that we specifically aim to model and predict, as well as the temporal dependence between two successive layers (year-months). Fishing gear was treated as a fixed effect. In combination with the offset in the linear predictor, the gear-specific intercept acts as an estimated scaling constant for the gear-specific relative efficiency per unit of effort. Variations across vessels were treated as a random effect consistent with what is done routinely when standardising for differences fishing effort (exposure time) by estimating catch per unit time (Wiff et al 2008). We refer to spatial layer as the realisation of the spatial process (i.e., map of count of juvenile cod) at one particular point in time. The final model proposed here is composed of 60 spatial layers, each one representing a month across the five years of analysis.

Table 2.1: Effects included in the final model.

	Definition	Description
	linear effects	
$z_{\alpha=1}$	Seasonal	pre-computed seasonal effect
$z_{\alpha=2}$	Depth	pre-computed depth effect
$z_{\alpha=3}$	Month	fixed effect of months
$z_{\alpha=4}$	Year	fixed effect of years
$z_{\alpha=5}$	fishing gear	each of nine fishing gears as fixed factors
	non-linear effects	
$c_{\gamma=1}$	spatio-temporal	structured spatial and year-month effect (Markov Random Field + AR1 temporal structure)
$c_{\gamma=2}$	Vessel	unstructured random effect

In an ideal situation, seasonal and depth effects should have been treated as structured non-linear effect in Eqn (2). Early attempts to fit this type of model to the data were unsuccessful because of the large number of the structured spatial layers that needed to be estimated which exceeded the limits of the available processing power (a computer with 3.5 GHz Intel Core and 16 GiB RAM was used). It seems reasonable to assume that the smaller-scale fish aggregations observed at sea are ephemeral structures and less likely to be repeatable across years, unless they are correlated with covariates such as depth. Assuming that seasonal spatial effects should typically be relatively large-scale processes and to make computation more tractable, we pre-computed spatial seasonal and depth effects from the whole set of data with spatial smoothing splines using GAMs. The intention here is to model seasonal and depth spatial features effectively but in a smaller dimension (at relatively broad spatial scales) using splines which are very computationally effective. Then, marginals (term used for the relative contribution of each effect) for seasonal and depth were taken from GAMs prediction and treated as fixed effect in model of Eqn (2). Thus, the linear effects for season ($z_{\alpha 1}$) and depth ($z_{\alpha 2}$) in Table 2.1 were pre-computed using the following GAM:

$$log[E(y_i|\mathbf{x})] = g(\mu) = \beta_0^G + \log(\beta_{of}^G) + ti(D) + ti(m,D) + te(x,y,m) + \ldots \quad (3)$$
$$+ te(x,y,by=m) + te(x,y) + Year + Gear$$

where β_0^G is the intercept and β_{0f}^G is log of trawling time treated as an offset. *ti* and *te* represent a tensor product smooth interactions and tensor product smooths applied over variables depth (*D*), month (*m*), longitude (*x*), latitude (*y*). *ti(D)* assumed a thin plate regression spline (*tp*). *ti(m)* assumed a cyclic cubic regression spline (*cc*). *ti(m,D)* used a *cc* and *tp* splines for months and depth, respectively. *te(x,y,m)* uses a thin plate spline with smoothing penalty (*ts*). *te(x,y)* and *te(x,y,by=m)* represent the random effect of the spatial component globally and by spatial month layers, respectively.

4.2.3.3 Structural Spatio-Temporal Analysis

The structural spatial correlation $c_{\gamma 1} = \boldsymbol{c} = \{c_1, \ldots, c_n\}$ was modelled using an intrinsic conditional autoregressive model, also known as the Besag model (Besag 1974). This model assumed that ICES grid cells that are adjacent in space show more similar number of undersized cod than areas that are not neighbours. A common assumption is to regard grid cells *i* and *j* as neighbours if they share a common border, denoted here as *i* ~ *j*. We denote the set of neighbours of region *i* and δ_i with size $n_{\delta i}$. The conditional distribution for c_i is:

$$c_i|\boldsymbol{c}_{-i}, \tau_c \sim N\left(\frac{1}{n_{\delta i}}\sum_{j\in\delta_i} c_j, \frac{1}{n_{\delta i}\tau_c}\right)$$

where τ_c is a precision parameter and $\boldsymbol{c}_{-i} = (c_1, \ldots, c_{i-1}, c_{i+1}, \ldots, c_n)^{\mathrm{T}}$. The joint distribution for **c** is:

$$\pi(\boldsymbol{c}|\tau_c) \propto exp\left(-\frac{\tau_c}{2}\sum_{i\sim j}\left(c_i - c_j\right)^2\right) \propto \left(-\frac{\tau_c}{2}\mathbf{c}^{\mathrm{T}}\mathbf{Q_s}\mathbf{c}\right)$$

where $\boldsymbol{Q}_s$ is the precision matrix of the spatial structured effect with entries:

$$Q_{i,j=}\begin{cases} n_{\delta i} & i=j, \\ -1 & i\sim j, \\ 0 & else. \end{cases}$$

The spatio-temporal interaction $c_\gamma(s,t)$ can be described as the product between purely spatial ($c_\gamma(s)$) and temporal basis function ($c_\gamma(t) = c_t$), here defined as first order autocorrelation process (AR(1)) between consecutive year-month layers:

$$c_t = \emptyset c_{t-1} + v_t$$

where observations are independent in time but spatially correlated:

$$v_t \sim N(0, \boldsymbol{Q}_s^{-1})$$

The spatio-temporal interaction was modelled as a Kronecker product (Blangiardo & Cameletti 2015) of a structured GMRF and first order autocorrelation in time, AR(1). The Kronecker product, here denoted by ⊗ , is an operation on two matrices representing the space and time into a partitioned matrix that has being broken into sections each one representing a point in time across all the space.

In this case, the precision matrix of the spatio-temporal interaction(**Q**) can be written as the Kronecker product between the precision of the temporal ($\boldsymbol{Q}_T$) and spatial ($\boldsymbol{Q}_s$) precision matrices as $\boldsymbol{Q} = \boldsymbol{Q_T} \otimes \boldsymbol{Q_s}$ (Blangiardo & Cameletti 2015). Here, $\boldsymbol{Q}_T$ is a tri-diagonal T-by-T matrix. The use of the Kronecker product allows the construction of a block matrix broken into space and time units. This is based on the assumption that each one of the *s* ICES grid cells (s_i), has an autoregressive structure on the time component *t*, which is independent of the ones in the other grid cells. The hyperparameter vector (θ) is defined as purely spatial (θ_1), temporal (θ_2) and the unstructured random effect for vessels (θ_3).

The final model contained 60 spatial layers (5 years by 12 months) in which each cell is connected with an AR(1) process. The dimensions of the parameter space for the volume of available data exceeded the limits of what could be processed in R-INLA with our space-time model as defined above, on a 3.5 GHz Intel Core and 16 GiB RAM. In order to make computation more tractable, we took advantage of the Bayesian approach, by splitting the data set into three blocks of 20 year-month layers each. This made fitting the model relatively tractable with a standard desktop computer. The model was fit to each block sequentially, using the posterior estimates of the model hyperparameters as priors for the following block of data.

4.2.4 Cross Validation

We used a cross-validation procedure in order to assess the ability of the model to forecast fish spatial distribution one-month ahead. The last 22 year-month layers were independently forecasted and the forecast compared with observed data. For each layer to be forecasted, priors for the model hyperparameters were taken as the final posteriors obtained from section 2.2.3 above, and latent random fields were estimated using the 12 months preceding the target layer. Following Blangiardo & Cameletti (2015), we use R^2 as a measurement of the goodness-of-fit of each year-month layer predicted. R^2 was computed as the squared of correlation between observations and prediction on log-log scale.

4.3 Results

Figure 2.1 shows the results of the marginal effect of depth across months using the pre-computed GAM model. Several months have a bi-modal distribution of juvenile cod in depth. This means that juveniles have preference for either shallow or deep waters. The overall performance of the model is adequate, with a R^2=0.92, indicating that explanatory variables chosen in Table 2.1 were adequate and resulting model explained a good percentage of the total variance (Figure 2.2). However, the model tends to overestimate the observed low count and zero observations, and thus the bottom-left corner in Figure 2.2 does not follow the 1:1 relationship between observations and predictions. The posterior predictive p-values agreed with this rather poor fitting of low observed counts. When observations matched prediction well, histograms of the predictive p-values will have a uniform distribution (Blangiardo & Cameletti 2015). However, in Figure 2.2 high frequency in p-values close to 0 and 1, is an indicator the model may not be adequate to represent both, very low or high numbers of undersized cod.

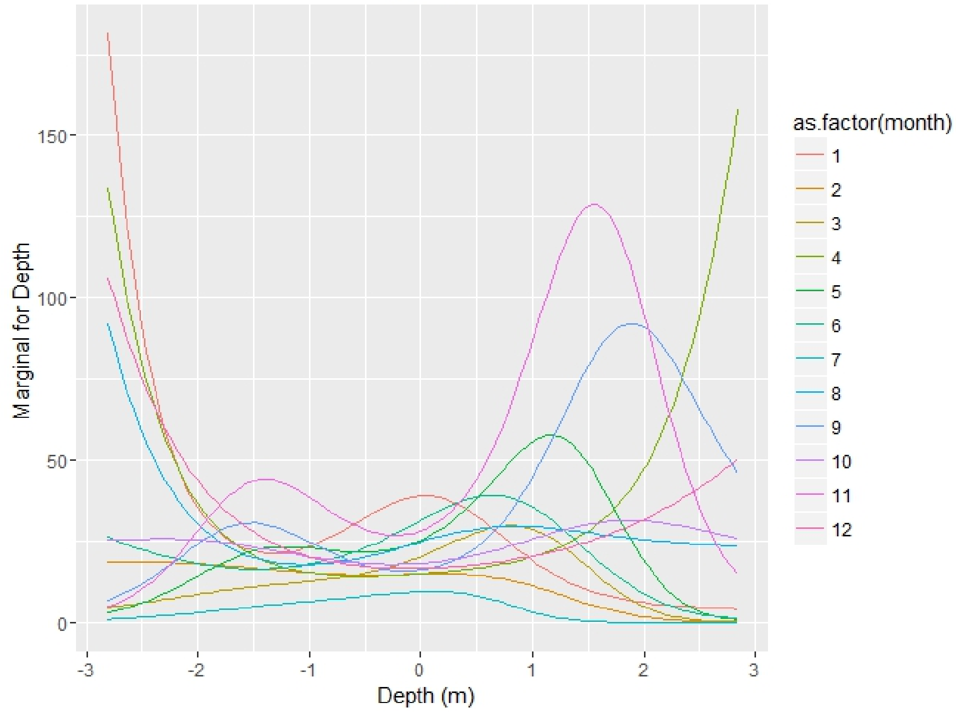

Figure 2.1: Marginals for depth (relative to the mean) by month from GAM model in Eqn (3).

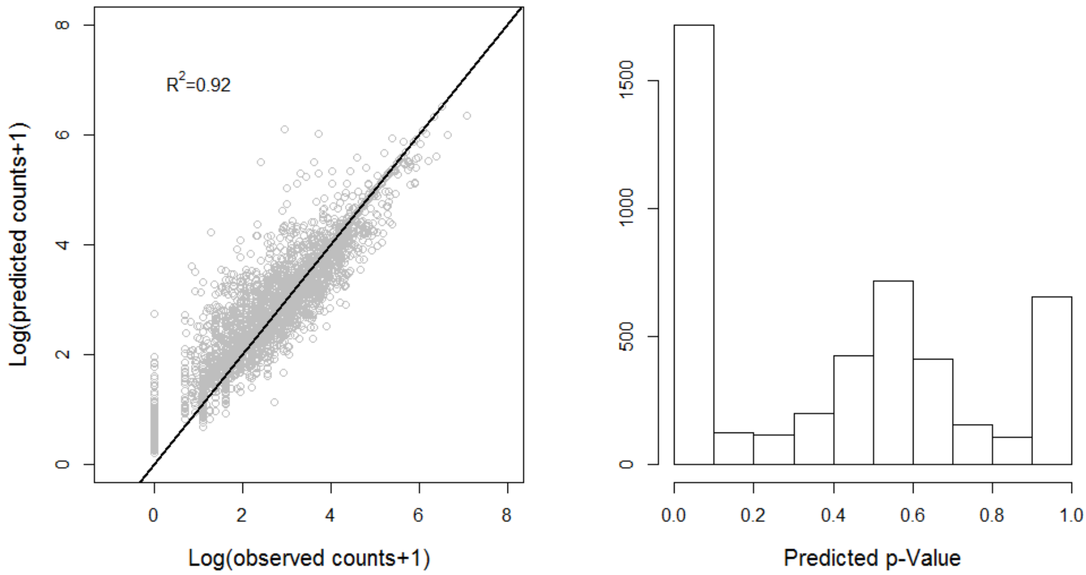

Figure 2.2: Goodness-of-fit of the general model in Eqn (2). Left panel shows the observation vs predictions in log scale. Right panel is the histogram of predicted p-values.

High seasonal and annual trends in spatial distribution were detected. In Figures 2.3 and 2.4 we show predictions for February and August across years, respectively. These two months were chosen because they represented a large number of observations coming from fisheries and survey data. Large values of observed CPUE in Figures 2.3 and 2.4 (large open circle), tend to be clustered also around large predicted values, indicating the model matched well spatial predictions. High intra-year variability is found (Figure 2.5). This variability seems to be in groups of three months (i.e the block January – March is relatively homogenous but different than blocks for April to June).

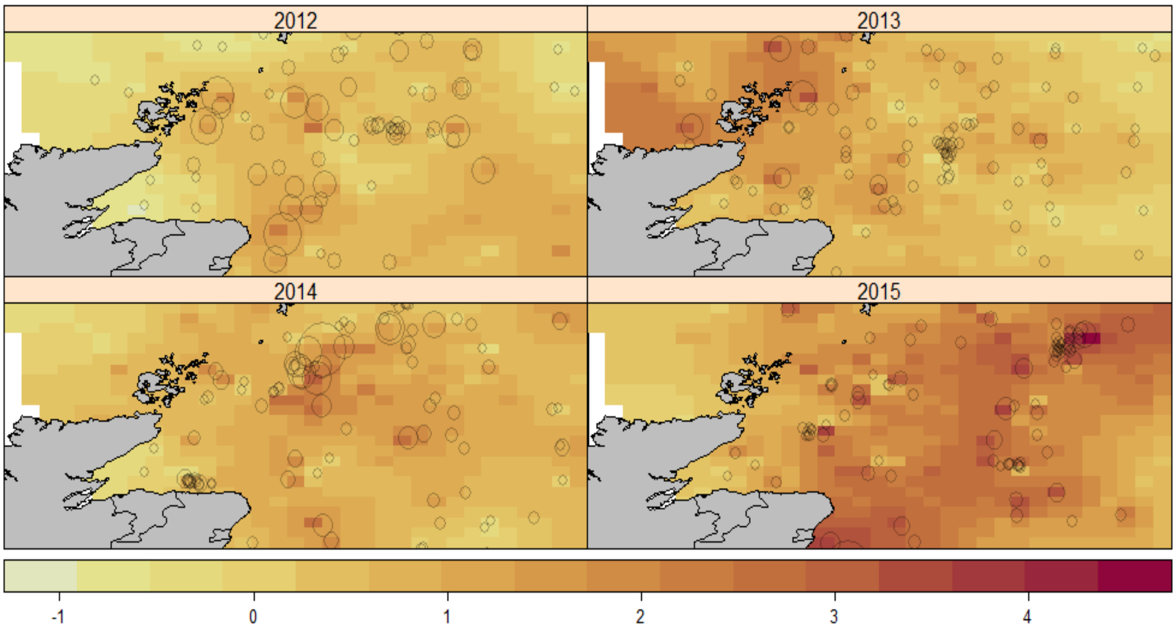

Figure 2.3: Predictions of juvenile cod (log of counts) on February between 2012 and 2015. Circles in each plot represent the observed juvenile cod CPUE from survey and observed data. Size of circles represent the relative magnitude of CPUE in each year-month.

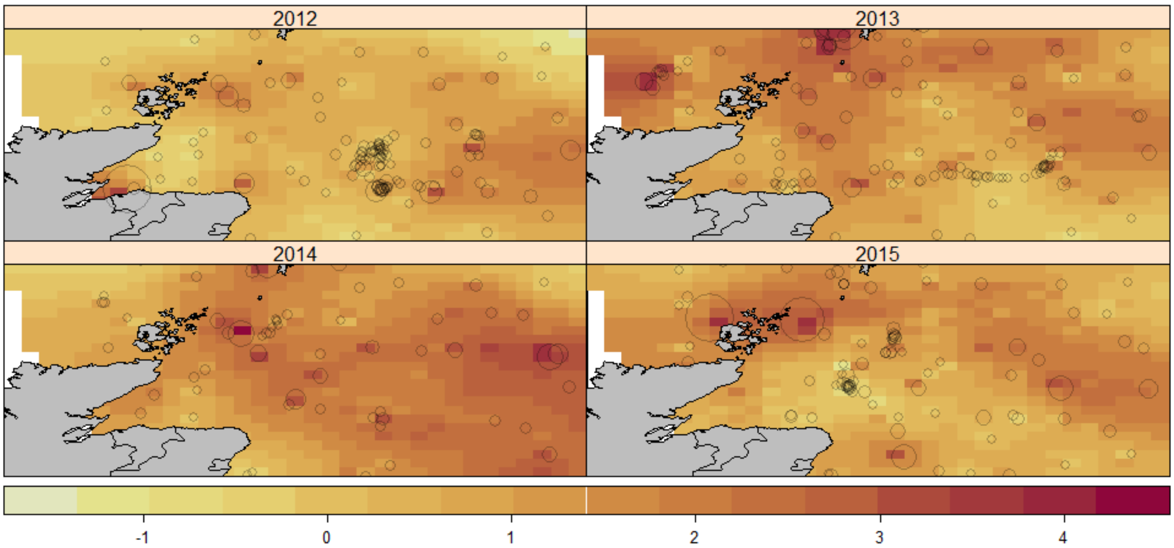

Figure 2.4: Predictions of juvenile cod (log of counts) on August between 2012 and 2015. Circles in each plot represent the observed juvenile cod CPUE from survey and observed data. Size of circles represent the relative magnitude of CPUE in each year-month.

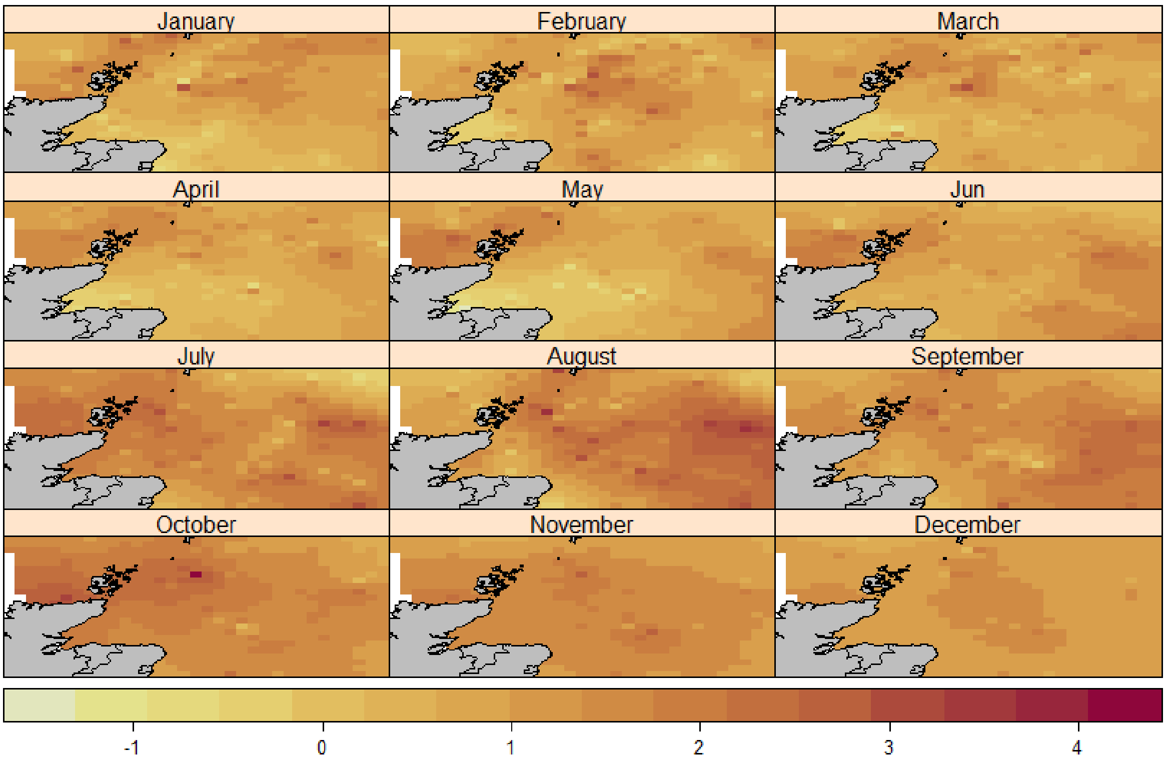

Figure 2.5: Predictions of juvenile cod (log of counts) by month for 2014.

Cross-validation shows high variability on the forecasting across months. R2 between forecasted and observed values ranged from 0.1 to 0.6 with an average R^2=0.35 across all months analysed (Figure 2.6). Forecasting in each year-month layer showed high variability (Figure 2.7), indicating the modelling framework was capable of reproducing the observed intra-annual variability.

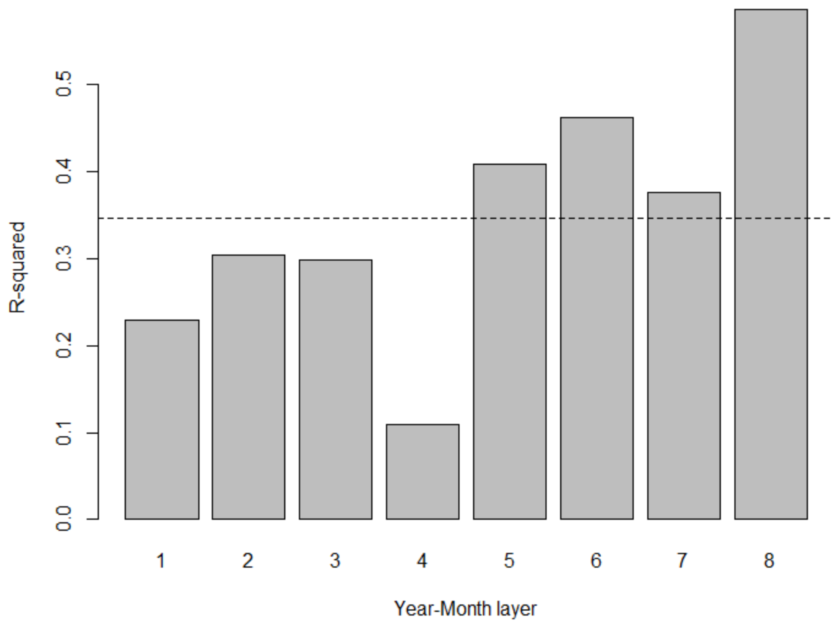

Figure 2.6: Goodness-of-fit of cross validation for each year-month layer predicted.

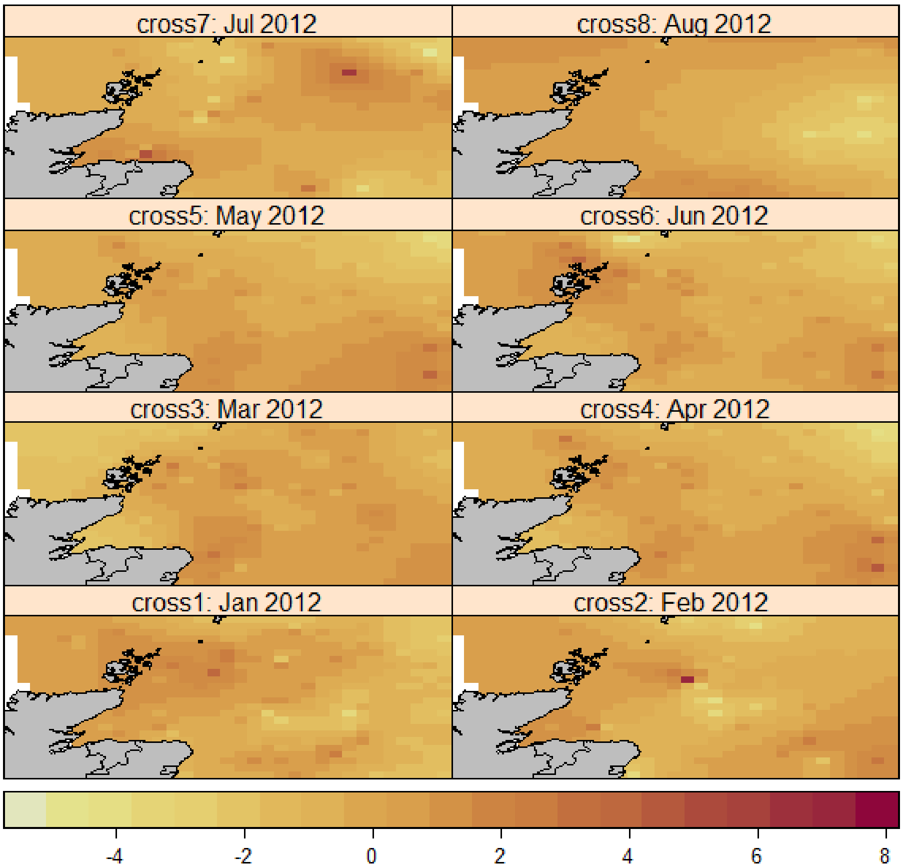

Figure 2.7: Forecasting of juvenile cod by months in relative units (scale to mean of each month). Results from cross-validation analysis, first block of years.

4.4 Discussion

The modelling approach merged commercial fishing and survey data to predict spatio-temporal distribution of juvenile cod in the North Sea. One of the main strengths of this approach is to combine information about catch rates of different vessels, gears, areas using the same idea underpinning effort standardisation. This means, using an additive model structure in which heterogeneity in catch rates caused by factors affecting catch rates are incorporated into a model via fixed or random effects. Another important strength of this approach is the use of GMRF, which allow the modelling of fine spatial resolution. In addition, Bayesian inference permits a comprehensive incorporation of the uncertainty, because it takes into account error in the observations and in the latent field. However, the use of GMRF is computationally intensive, and thus, some trade-off had to be considered. We assume that seasonal variations in spatial distribution of juvenile cod occurs at a wide spatial resolution, and thus, we applied a spline based model for those fixed effects. There are a few ways to improve this issue. One solution is to model seasonal spatial structure using also GMRF, but computation is not tractable for the spatial scale considered using a state of art desktop computer. The overall model could become tractable if the spatial scale became smaller. This means, divided the modelled area of the North Sea into smaller units that can be assumed as homogenous in some juvenile cod characteristics (e.g. distribution of nursery areas). Nevertheless, the model seems to represent well the fine spatio-temporal distribution of hotspots of juvenile cod at reasonable computation cost, and this will be an important framework for developing effective real-time management schemes in the future. The modelling of a fine scale spatial pattern in many cases will prove more useful for the fishing fleet in terms of fishing suitability maps. Such approaches would be interesting when modelling species showing strong aggregating behaviour, or species that are tightly linked to a certain habitat (see Lichstein et al. 2002~~).~~).

The real-time spatial management systems in place in some US Pacific fisheries are based on almost daily submissions of logbook catch-data by the skippers (see Little et al. 2015). In such systems, data rarely feed into a predictive model, but instead into algorithms creating polygon-based by-catch risk zones, much like the Real Time Closures operating in the North Sea mixed demersal fisheries. Within such framework, the real-time aspect of data processing became a key issue, because the empirical nature of the approach does not allow predictions. As an alternative, a model-based approach produces predictions of risk or fishing suitability maps with commercial fishing operations data continuously feeding a model upon which model predictions can be updated. Modelling framework proposed here, can be seen as a medium complexity alternative between the pure empirical approach of polygon-based bycatch analysis and the mechanistic approach proposed in Kristensen et al (2015). Such mechanistic approach is based on survey data alone, in which population parameters, such as growth and fishing mortality, were incorporated into a spatio-temporal statistical model to estimate how size and spatial distributions developed in time, which could predict by-catch risks for any location and time. However, there is a trade-off between increasing ecological realism and computational feasibility, and acquiring enough data to ensure reliable estimated parameters. We do not offer a mechanistic inference for the underlying process determining the observed pattern, but rather we modelled observations to predict the spatio-temporal dynamics of juvenile cod. The pure mechanistic approach in Kristensen et al (2015) relies on survey-based information, while the framework proposed here added also commercial fishing operations.

There are many other sources of commercial data that could potentially be used for inference about spatial distribution of juvenile cod. In this paper, data from the Scottish discard observation program were used, but a natural extension could be to use landings data, which are available on much higher temporal resolution as all landing have to reported, while only a small subset of fishing trips have on-board observers. However, the challenges with using landings data to infer spatial distribution lie in how to properly estimate the total catch at each haul location, as only the landed component is recorded in that data, and geo-referencing is not available on a haul-by-haul resolution. Under a Landings Obligations in the EU, this may well change as the entire catch must be landed (initially TAC species), which could open up this under-utilized source of catch data for use in spatio-temporal by-catch models.

The approach proposed here, in which commercial data is coupled with scientific survey-data, recognises the strengths of each data source; the high temporal resolution of commercial data and high spatial resolution of the survey data. The results presented a modelling approach for further utilizing non-standard fisheries data to infer spatio-temporal distribution. This is a highly topical area for fisheries research today, as there is an increasing demand to develop high-resolution spatio-temporal modelling approaches to infer fish-stock distribution for management purposes.

5 How Spatio-Temporal Information is Used in the Management of a Pacific Groundfish Fishery

The Pollock stock in the EBS is considered to be one of the best managed fisheries in the world. In 2015 it comprised 67% of the total groundfish catch off of Alaska (Ianelli et al. 2015). The status of the Pollock stock is evaluated annually with respect to two components of performance: a) stock status with regard to fishing mortality and yield; and b) bycatch corresponding to the incidental catch of chum and chinook salmon. Thus, bycatch targets form explicit management targets that are supported by federal legislation. Exceeding these targets would result in industry shut-down. This legislation effectively incentivises the fleet to develop optimal targeting strategies.

To meet these strict targets for bycatch reduction the industry began developing methods for the real-time, verifiable reporting of bycatch at sea in the 1990s. Measures included the use of a trained observers on the main fleet segments, access to vessel-specific VMS data, and triggers for time and area closures. In combination these measures allowed so-called “rolling hotspots” (RHS) closures to be identified, communicated and monitored.

5.1 Reducing Salmon Bycatch in the Pollock Fishery Using Rolling Hotspots

Salmon (chum and chinook) bycatch has been a management concern in the Pollock fishery for over twenty years. In the 1990s, regulatory Salmon Saving Area Plans (SSAPs) were introduced to manage salmon discards. However, with experience it became evident that SSAPs sometimes shifted vessels onto fishing grounds with even higher bycatch rates. In January 1999, the Pollock Conservation Cooperative (PCC) and the High Seas Catchers' Cooperative (HSCC) signed an intercooperative agreement (ICA) to jointly harvest and allow the transfer of Pollock quota between cooperative members. Although it is not written into US fisheries legislation, the ICA is a private and contractually binding agreement for all cooperative members. Individual vessels belonging to the PCC and HSCC are bound by the conditions of the ICA and compliance is monitored by the cooperatives themselves. In 2001, the ICA was expanded to include catcher vessels delivering to motherships and shoreside processing plants. The ICA introduced voluntary RHS closures to better respond to the rapidity of chum salmon spatial and temporal dynamics. RHS closures are temporary area closures which may be fished depending on the particular cooperative's (or vessel in the case of chinook salmon) bycatch performance. RHS were introduced for chum and chinook salmon in 2001 and 2002, respectively. In October 2005, under Amendment 84 to the Bering Sea and Aleutian Islands Fishery Management Plan (BSAI FMP), all vessels participating in the RHS programme were made exempt from regulatory SSAPs.

Sea State, Inc. is contracted by the PCC to receive, monitor and evaluate catch and bycatch data in the chum and chinook salmon RHS programme on behalf of the cooperative. The locations of the RHS are determined from real-time bycatch information provided by trained observers in combination with landing reports from shoreside processors. Under the RHS scheme vessels must report bycatch information within 24-48 h. Sea State, Inc. analyses bycatch data to identify straight-edged polygon hotspots on a bi-weekly basis. Each week fishing cooperatives are allocated to one of three "tiers" according to their bycatch performance for chum salmon over the last three weeks, expressed relative to the fleet's average performance. For chinook salmon the comparisons are made over the same time periods at the level of the individual vessel, not at the cooperative level. Tier 1 is for cooperatives/vessels with the best bycatch performance, tier 2 for medium performance cooperatives/vessels and tier 3 for worst performing cooperatives/vessels. These tiers determine the amount of time a cooperative/vessel will be forbidden to fish for Pollock within the RHS closures the following week. Tier 3 vessels are prohibited from RHS for 7 days, tier 2 vessel for 3 days, while the tier 1 vessels are not prohibited from fishing in the RHS closures. In other words, cooperatives/vessels that are able to fish cleanly are not restricted from the RHS closures. This form of performance-defined access incentivises the need for a cooperative or vessel to reduce bycatch.

To support the implementation of RHS closures, salmon bycatch data are submitted to the authorities in quasi-real-time (within 24-48 hours) by observers on-board the vessels. There is 100% fishery observer coverage for all vessels in the Pollock fishery that is entirely industry funded. The bycatch data are uploaded to a central National Oceanic and Atmospheric Administration database but each cooperative has access to their own data. To support the ICA, Sea State, Inc. monitors the compliance of individual vessels belonging to the cooperative with the terms of the RHS scheme by comparing VMS records against observer data. Sea State Inc.'s compliance monitoring is audited to avoid any criticism of a third-party monitoring. According to the American Fisheries Act (AFA), the ICA must report their bycatch performance annually (PCC and HSCC 2013) because in essence the ICA is co-managing the stock with the North Pacific Fisheries Management Council (NPFMC; http://www.npfmc.org/). Sea State Inc. reports on the performance of individual vessels in terms of bycatch to the ICA on a weekly, rolling fortnightly and seasonal basis. As well, reports are prepared for a range of public agencies such as tribal stakeholders.

5.2 Technicalities of Implementing a Bycatch Reduction Scheme for Chinook Salmon

The chinook salmon is the largest salmon species in the Pacific and is known as the "king salmon" in Alaska. The ICA developed an incentive plan agreement (IPA) and performance-standard requirements specifically to minimize bycatch of chinook salmon so as to meet these regulatory requirements (failure to meet them would result in industry tie-up). The IPA was informed by the experience gained in the development and refinement of RHS programs. It restricts the Pollock fishing opportunities of vessels with poor chinook bycatch performance while allowing vessels with good performance unimpeded access to the fishing grounds (tier 1, 2, 3 designation). Avoiding such restrictions reduces operating costs and allows for the production of high-valued products, thus increasing profitability. The IPA rewards good vessel chinook bycatch performance irrespective of Pollock or chinook salmon abundance. The chinook IPA can be found online at https://alaskafisheries.noaa.gov/sites/default/files/chinook_salmon_ipa_2010.pdf.

Sharing information about chinook bycatch is at the heart of the IPA. The specific components of the IPA for chinook therefore include: (1) data gathering, bycatch monitoring, reporting, and information sharing; (2) identification of "bycatch avoidance areas" which are the equivalent of RHS; and (3) fishing-area prohibitions for vessels with poor bycatch performance. In this way, the IPA operationally defines an approach to achieving the bycatch regulations of the BSAI FMP that was designed by industry using past experience with spatial selectivity.

At the start of the 2011 fishery, Amendment 91 to BSAI FMP came into effect. Amendment 91 is an innovative approach to managing chinook bycatch that combines a prohibited species catch (PSC) limit, or cap, on the amount of chinook salmon that may be caught incidentally by the fishery and performance-standard requirements designed to minimize bycatch to the extent practicable in all years. The total chinook salmon PSC cap of 60,000, with a performance standard, or target, of 47,591 chinook was incorporated into the IPA.

5.2.1 Effectiveness of the IPA for Chinook Salmon

The success of the bycatch reduction programme can be difficult to demonstrate as the most dramatic reductions occurred by keeping the majority of the fleet from moving into areas of very high bycatch. This is analogous to trying to prove the size of a fish that got away or demonstrating prevention of an accident that never happened. Logically, the potential for further high bycatch exists when one or two vessels encounter extreme hotspots. Without disseminating knowledge of the location of these hotspots other boats would have likely entered those areas.

A recent report assessed the impact of Amendment 91 in 2011 for reducing chinook bycatch on the bycatch performance of the fleet and individual vessels (Madsen and Haflinger 2015). Chinook bycatch performance (number of chinook per ton of pollock caught) of IPA vessels improved following the implementation of Amendment 91 (Fig. 3.1), as compared with the previous four years. Variability in environmental conditions and salmon abundance could have played a role throughout this time period, however, the improvement is consistent over time. Vessel bycatch rates are currently among the lowest on record, and the variability of bycatch rates among vessels has been reduced relative to pre-2011 years with similar average bycatch rates. This provides quantitative evidence of the effectiveness of the vessel-level, by-catch incentives.

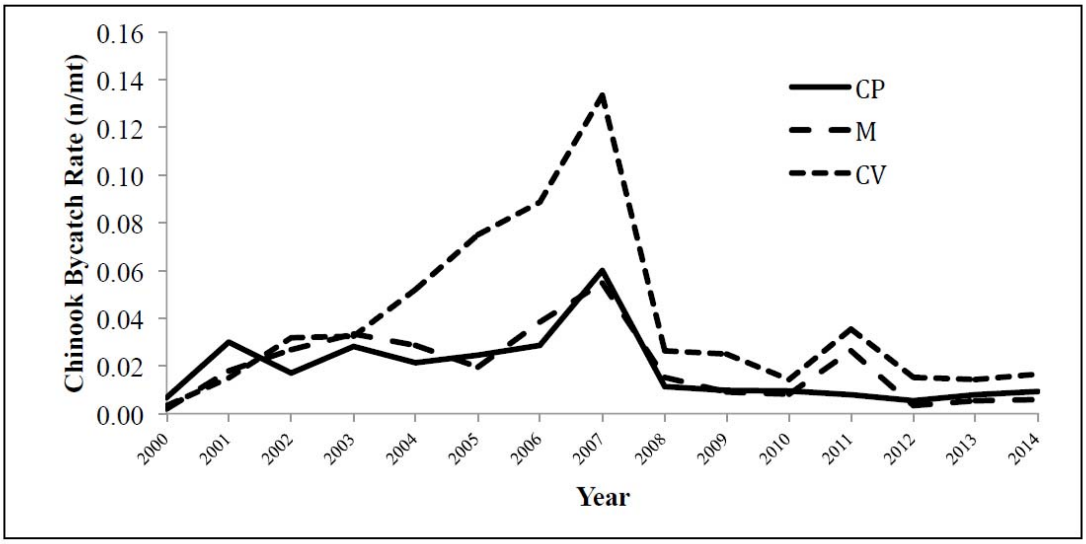

Figure 3.1: Chinook bycatch rates (n/mt) by year for the Catcher Processor (CP), Catcher Vessel (CV), and Mothership (M) pollock fishing sectors in the Bering Sea (from Madsen and Haflinger 2015).

5.3 The Evolving Role of Sea State Inc. in Monitoring Bycatch

When the bycatch reduction programme began in late 1990s (with the first ICA) the industry began funding the collection of observer data. Early attempts to reduce bycatch of salmon through sharing observer data were largely ineffective due to the "race for fish". However, the Pollock industry continued to fund the observer program as part of the arrangement by which they access the resource. In this respect, the catch data is a useful byproduct of the fishery. At this time Sea State, Inc. was contracted to manage the database. To do this, they pioneered early forms of information sharing and mapping capabilities by equipping wheelhouses with plotters. Overall, these reporting procedures were felt to be successful in that fishing seasons were kept open longer than would otherwise have been expected, due to cooperation amongst fleet in moving away from areas of high bycatch.

In 1998 the AFA effectively ended the "race for fish" in the Pollock sector. In 1999 all sectors came together to write the first ICA designed to prevent closures of their fishery based on chum and chinook salmon bycatch. This effectively restructured the industry into a fully functioning cooperative and ushered in greater engagement in co-management of the resource. The regulatory need to meet strict bycatch targets under full observer coverage has enshrined spatial selectivity as an operating principle. Over the years Sea State, Inc. has incorporated a range of information and communication technology (ICT) advances (summarised below). Spatial management measures also include multiple conference calls that often include skippers as well as cooperative managers having a fleet-wide knowledge of cooperative performance. In combination, these methods effectively constrain the behaviour of individual vessels belonging to cooperatives operating under the highly prescriptive terms of the IPA.

5.4 Summary of discussions with Sea State Inc.

The one-day discussion was split into three sections: 1) general observations on bycatch reduction solicited in an unstructured way; 2) discussions focussed around a demonstration of the software used by Sea State Inc. to gather data, analyse data, report data and disseminate information to the industry; and 3) a discussion of modelling approaches that could be used to enhance the analysis of available industry data. This write-up summarises those discussions by focusing principally on the following components of the IPA: data gathering, reporting, bycatch monitoring and information sharing (Sections 3.4.2, 3.4.3, 3.4.4 and 3.4.5, respectively). General observations that fall outside those components are noted immediately below.

5.4.1 General Observations about Bycatch Reduction Schemes

- The IPA can be considered a direct example of results-based management (Nielsen et al. 2015) specific to achieving federal regulations related to bycatch reduction. The IPA defines the means by which the ICA effectively co-manages the stock with the NPFMC.
- Bycatch reduction schemes work well when the fishery is not open access and that there is no “race for fish”. It would be impossible to implement them in an open access fishery. As was described above, the “race for fish” in the EBS effectively ended in 1999 with the implementation of the AFA and the restructuring of the industry into fishing cooperatives through the ICA.
- The IPA is in principle voluntary but only in the sense that membership in a given cooperative is voluntary. Once a vessel joins the cooperative then it is bound to the terms and condition of the cooperative’s own IPA. In this sense, the cooperative is ensuring compliance on vessels. The regulatory agency is uninvolved in the technical means by which bycatch is reduced, i.e., the IPA.
- The ICA uses a “name and shame” approach to identifying individual vessels with poor bycatch performance. This works by applying a form of moral persuasion for skippers to comply with industry-defined bycatch targets. The reporting performance of an individual vessel confers bragging rights on

vessels that are performing well. There is a distinct skipper effect on bycatch performance.

- Bycatch of chinook and chum salmon is approximately 4% and 1% of the total stock size of each salmon stock, respectively. Because bycatch is such a small proportion of the total mortality on salmon, bycatch reduction will have minimal impact on overall rates of stock rebuilding. The creation of bycatch targets by Amendment 91 of the BSAI FMP on the Pollock fishery can be more accurately viewed as a mechanism by which the Alaskan fishing industry shares the burden of salmon conservation.

5.4.2 Data Gathering

As noted above, the observer programme provides 100% coverage and is financed exclusively by industry. Private contractors supply personnel and the National Marine Fisheries Service (NMFS) providing the observer training including software training. NMFS uses the observer data to monitor bycatch performance of the fleet. Observers have differing degrees of sophistication in how they interact with the software. At sea, each haul is sampled for species composition and the total weight caught is recorded. Large hauls are handled differently with some sub-sampling and then raised to the proportion of total catch which was sub-sampled.

Observer data from catcher-processors and larger trawl catcher vessels are generally available to Sea State Inc. one to two days after a haul has come aboard a vessel. The magnitude of the time lag depends on how often the observers enter their observations into the computers they use for reporting and how often they actually send those data electronically to the observer office at NMFS. Once the data arrives at NMFS it is scanned for obvious errors by software prior to being made available for download by users (e.g., government scientists, Sea State Inc on behalf of the cooperative). There are often changes to the data that emerge up to several months after initial receipt, as observers make their way back to shore for a final "debriefing" after which the data are considered final.

The process of downloading data has been automated by Sea State Inc. and the program currently runs four times daily. In the past, Sea State Inc. used to "hand-build" reports with various tables of bycatch rates and the types of maps that a particular bycatch avoidance program required. However, they are moving towards automating as much of the reporting as possible so that information can be provided to the fleet in a more timely way (independently of human analysts to process and disseminate). The email message shown in Fig. 3.2a is one example. It was generated at 02:00, after the midnight download of data from NMFS was processed and analyzed by the software to see if any alarms had been triggered. As it did trigger an alarm the email was distributed to the fleet automatically.

Shoreside landings data are also an important component of the bycatch database. These landings data cannot be accessed until after a vessel actually offloads. Shoreside plant personnel generally send an initial report to the State of Alaska within 12 hours of a vessel landing, and generally have a fully updated and edited report with final numbers, within several days.

5.4.3 Reporting

Karl Haflinger initially created the software for capturing, analysing and distributing data via a protected website. The software was upgraded several years ago by a programmer in Juneau (Eric Torgerson, Chordata Inc) with expertise in fisheries software and database development. The EBS fishing industry has excellent access to the internet at sea via satellite communications. The software is designed to make efficient use of bandwidth by prioritising the transfer of essential information. The Sea State Inc. databases are physically housed in a server located in the states of Michigan and Arizona which has thus far proven to be reliable, secure and cost-effective.

Although the Pollock fishery is the focus of this FISA report, the Sea State Inc. database covers all of the federal groundfish fisheries in the EBS and Aleutian Island region, and the whiting fishery off the coast of the Pacific Northwest (Washington and Oregon). Several different data types are stored for analysis by the software including: observer data; VMS data; production data from shoreside plants and catcher/processors; and shoreside landings data. The observer data are obtained on a semi-continuous basis (as described above). Different types of data products are relevant to different fisheries so there is a degree of customised programming that needs to be done for each reporting requirement.

A key principle in software design is that individual vessels can access their own data and use the Sea State Inc. software to generate reports that summarise their bycatch caught to date against the allocation they have for a given fishing year. They can schedule the delivery (via email) of reports according to how frequently they wish to be updated. It is important to note that an individual vessel does not have access to catch information about target species from other vessels although catch information about bycatch species is shared across the fleet. Many fishing cooperatives have cooperative managers who have access to these information sources and who would undertake some of the detailed analyses of performance. Karl Haflinger does this analysis in his role as cooperative manager for several smaller fleets. Larger fleets would have their own analyst, e.g. the cooperative manager.

5.4.4 Bycatch Monitoring

Bycatch rates are reported in a standardised unit that corresponds to the number or weight of salmon caught incidentally divided by the metric tonnes of Pollock caught. Further details of this calculation can be found in the IPA https://alaskafisheries.noaa.gov/sustainablefisheries/bycatch/salmon/chinook/ipa/chinook_salmon_ipa_2010.pdf

If a haul having high bycatch is reported then an alert is sent to the industry with a link to a report showing a map with the geographic coordinates of the haul and basic information about the bycatch (Figure 3.2a).

```
This message was generated on 9/21/2015 at 2:10 AM

All high bycatch hauls:
https://acct.seastateinc.com/Seastate/Members/AfaPollockMap.aspx

AFA CP Haul 159 on 9/20/2015 has a total catch of 112 Chum Salmon.
This is 1.1 x the alarm threshold of 100.
Latitude: 55 8.90 N
Longitude: 167 32.00 W
VMS track:
https://acct.seastateinc.com/Seastate/a.aspx?p=1&a=1799&h=6ef5ff0ba571b21a81882b50969b8c49
```

Figure 3.2a: Example of an alert sent automatically by email to cooperative members to report a haul having a high bycatch. Haul VMS track is given on a clickable link. This shows high bycatch tracks from the last 2 weeks of the fishery. Skippers can access this information at sea by reading their email.

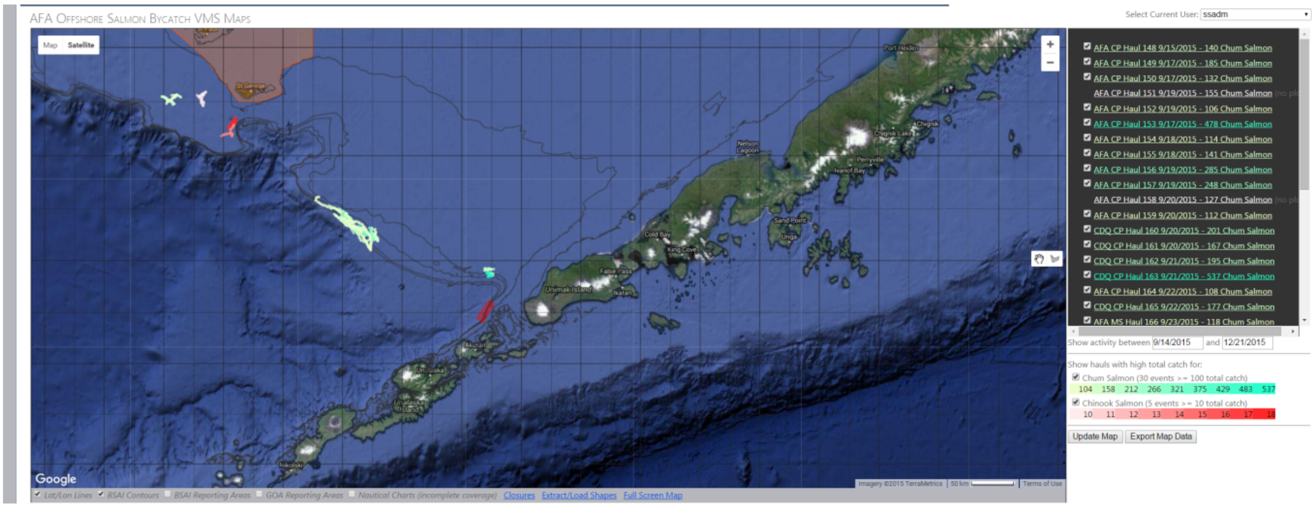

Figure 3.2b: Example of the map that is embedded in the alert message (Fig 3.2a). The VMS tracks of hauls that caught chum salmon are shown in blue and chinook in red. The darker the colour the higher the bycatch rates (scales are shown at the bottom right). The individual hauls that are shown are identified on the top right of the image.

5.4.5 Information Sharing

A key design principle in designing the IPA has been that individual skippers should be able to access their own data on demand. Different levels of access are accorded to different roles: for example, the cooperative manager can access data for all vessels in the cooperative. Over time, the skippers in the Pollock fisheries have become “information junkies” in the sense that there is a very high demand for the type of highly resolved spatio-temporal information that is currently being disseminated. The increasing reliance on the information being generated by the observer programme serves to reinforce industry’s commitment to funding the programme. Fishing success, including profitability, is increasingly determined by having successful tactics. Tactical fishing strategies (where to fish and when) are being heavily influenced by real-time reporting. This incentivisation to avoid bycatch is partly driven by the terms of the IPA which grants access to fishing grounds (the RHSs) according to bycatch performance. In a sense this creates a virtuous circle within the cooperative: improved information leads to improved profitability leads to improved information.

5.4.6 Quantitative Analysis of Bycatch Data

5.4.6.1 Rolling Hotspots and Access to Them

Karl Haflinger, serving as either the administrator (for large fisheries such as EBS Pollock) or cooperative manager (for some smaller fisheries in the EBS), has access to all data for the cooperative. Using the bycatch data he designs the RHS as polygons according to pre-specified designs (with some flexibility). This requires some expert judgement, incorporating experience of spatial and temporal patterns in distribution of the bycatch species. Some RHS closures repeat themselves over time and there may be short-term impacts of temperature (warm vs. cold years) and longer-term climate change impacts. Week-old data is the most relevant for defining RHS. Data that is more than two weeks old is starting to get old, illustrating how quickly the information value of bycatch data decays. This rapid decay is relevant to developing more advanced data processing (i.e., described in Section 2). Models need to run on highly resolved time and space scales to be useful to the industry.

RHS closures are in effect until the next closures are announced in 3 or 4 day intervals, when a RHS closure could be discontinued if data from Tier 1 vessels showed no problems. In the absence of new data from the closed area a RHS closure may simply be extended for up to two weeks. In this case Sea State will simply re-announce the same closure coordinates at the next scheduled closure announcement. As noted above, Tier 3 vessels are prohibited from RHS for 7 days, tier 2 vessels for 3 days, while the tier 1 vessel are not prohibited from fishing in the RHS closures. In the case of chum salmon, Tier 1 cooperatives are defined as those cooperatives having in aggregate (over the previous 3 weeks) less than 75% of the 2-week average bycatch rate. Tier 2 cooperatives are those with an aggregate catch of chum salmon between 75% and 125% of the average rate. Tier 3 cooperatives have greater than 125% of the average rate. For chinook salmon, which is the more rare and sought-after species, these same definitions apply but at the individual vessel level instead of at the level of the cooperative. Individual vessels frequently have their own tradable quota of chinook.

5.4.6.2 Spatial Mapping

The software Sea State Inc. uses has visually impressive, GIS-like mapping capabilities (see Fig. 3.2b, Fig 3.3). The bycatch information has been reported in the form of “heat maps” with catch rates reported for grid cell. Alternatively, the information can be presented by showing tracks for the individual hauls with the track-specific bycatch information revealed when the mouse is positioned on the track. Bathymetry is shown for reference. The programming utilises the functionality of Google Earth and is therefore non-proprietorial.

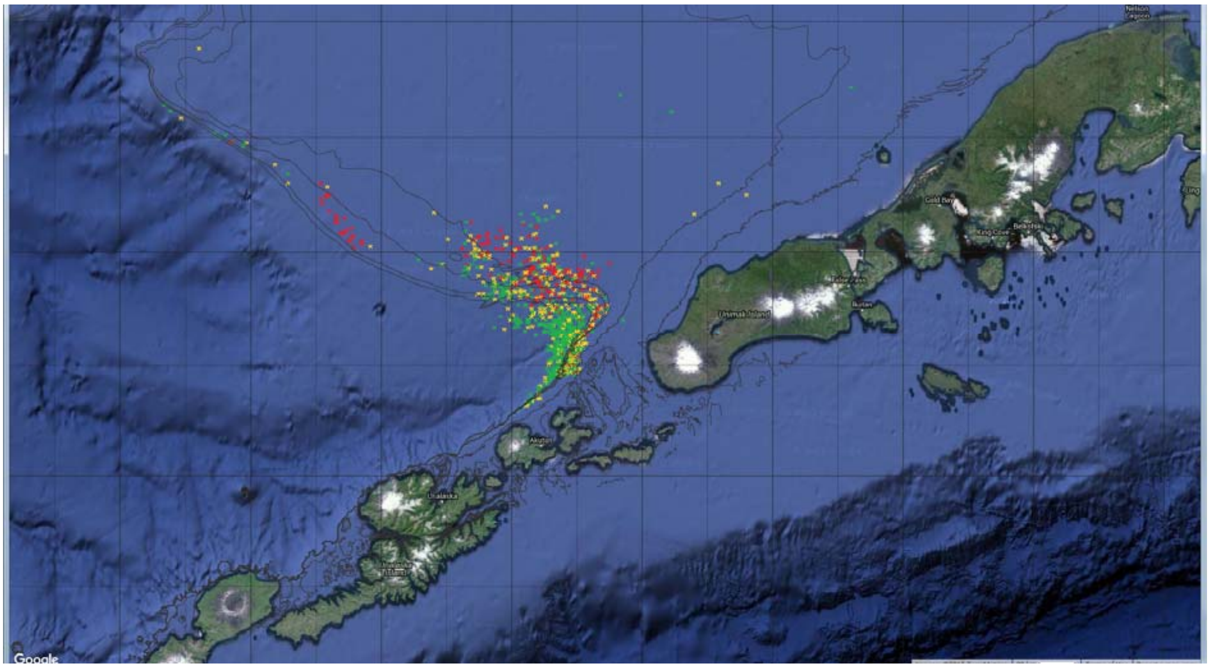

Fig. 3.3: Map indicating the lowest 50% of salmon bycatch values (green), 50-80% values (yellow), and top 20% (red). Additional layers can be added to show the locations of rolling hotspots closures.

5.4.6.3 Future Opportunities for Modelling

Sea State Inc. is currently using a model-free approach to the treatment of data in the sense there is no attempt to fit statistical models to interpolate information across time or space scales or identify hotspots via analytical means. This overlooks the obvious information content of having repeated measures of abundance generated from multiple hauls in the same area. This also overlooks the large amount of information contained in past data, which can serve to map background expectation of catch (long-term average) where up-to-date data are unavailable. Both the smoothing of real time data and the blending of real-time and historical data require specifically designed models to apply appropriate weighting to each data source. Now that the operational aspects of gathering, reporting and disseminating information about bycatch have been addressed there is an opportunity to consider opportunities for modelling the data for scientific purposes (e.g., linking salmon distribution to oceanographic features) or industry use (e.g., predicting where hotspots are most likely to be at any given time, similar to weather forecasting). The following points were made during discussion of the research potential of the bycatch data.

- Research use of the data is made somewhat problematic by issues of confidentiality. There are ways around this: either by anonymising the individual vessel data or by aggregation across defined sectors or sub-units. The following research ideas were identified by our discussions.
- Other analytical issues include how to distribute bycatch abundance for a single haul over the spatial path of that haul considering that bycatch abundance is not uniformly distributed.
- There are types of data that would be very useful to have for modelling that are not contained in the dataset. Temperature is the most obvious explanatory variable to consider. The industry does not have the ability to generate temperature data in a scientifically valid way (e.g., calibration issues). However, temperature data are available from other sources and could probably be integrated with the bycatch or catch datasets. Frontal structure, positions of eddies and also tides are all relevant to location of fish. Over longer time scales this could be pursued through targeted research programme.
- In essence, the bycatch database is showing the industry where bycatch performance is bad (negative result) but not necessarily all of the places where bycatch performance may be good (positive result). This bias highlights risk of bycatch: the “hotspots” but it is not showing all of the “coldspots”. More could be done to provide this sort of information through modelling.
- Spatio-temporal hotspot models can be used to assess the co-occurrence of a target and a non-target species to achieve management goals relevant to the threatened species (Ward et al 2015).

5.4.7 Future Directions

Sea State Inc. is looking to provide more automated hotspot maps via email to vessels that cannot browse their website due to bandwidth considerations. In future, they will hopefully be more sophisticated than the "point" map shown above (Fig. 5.3). For example, geostatistical methods could be applied to insure that bycatch trends have some statistical significance and can be relied on by fishermen.

The observer programme is the backbone of the database, however, these data are expensive to collect. In addition to salaries there are considerable travel expense incurred by flying observers to and from ports in remote areas of Alaska. Electronic monitoring (EM) by CCTV is one means of reducing these costs and it is being trialled in the whiting fisheries off coast of Washington and Oregon. These are full-retention fisheries, so all catch is essentially counted at the time of delivery to either a plant or mothership. The aim of EM is to insure that vessels are complying with the no-discard rules. Sea State and Chordata are working currently on a project to reduce the time spent reviewing CCTV data by developing computer algorithms to flag activity on a fishing vessel that may indicate when discards could potentially occur.

5.5 Relevance of EBS Fisheries to Scotland

Fisheries on the west coast of North America have developed comparatively recently and consequently have evolved very differently to the traditional fisheries on the east coast of North America which were developed like European fisheries (Little et al. 2014). West coast fisheries therefore have more examples of innovative styles of co-management (e.g., application of individual transferrable quotas, industry funded observer programmes). It could be argued that they are closer to results-based management than European fisheries. In that respect, they serve as an interesting model for comparison. More specific similarities and differences are identified below (the list is not intended to be comprehensive).

5.5.1 Similarities

Reducing salmon bycatch was established as a regulatory requirement for the Pollock fishery in response to societal concerns about conservation of Alaskan salmon. This is broadly similar to the concern about discarding rates at sea globally and in Europe.

Since 2007, Scotland has operated a voluntary system of “real time” closures (RTCs) to help the continuing recovery of cod stocks in the waters around Scotland (Needle and Catarino 2011). RTCs are one of several ‘conservation credit’ measures that have been taken under the EU’s Cod Recovery Plan. Scottish vessels are allowed more time at sea in return for adopting conservation-minded fishing practices including observing RTCs. These RTCs are analogous to the RHS in the Eastern Bering Sea used to avoid salmon bycatch.

The Scottish demersal fleet is organised into separate producer organisations (POs) and there are several clear parallels which can be drawn between POs and fishing cooperatives on the west coast. In both cases, national quota is allocated to the PO or cooperative by the regulator and then distributed by the PO or cooperative to individual vessels. Both POs and cooperatives have reserve quota that can be drawn on by skippers who have exhausted their vessel allocation (who may also be free to lease quota separately). PO and cooperative managers have access to landings data which allows them to review how much quota remains and facilitate transfers. On the west coast these advantages of membership in a cooperative come with the obligation to fulfil the cooperatives targets for bycatch reduction as specified by the IPA (see Section 3.3). Scottish POs have the e-logbooks which are in several respects better real-time databases given that landings information about quota species is uploaded within 2 hours of the haul coming on-board.

Both the Alaskan and Scottish fishing industries have access to high quality, real-time information. In the EBS fisheries observer data are generally available to Sea State Inc. one to two days after a haul has come onboard. As there is no discarding in these fisheries, the catch matches the landings data. In Scotland the current version of the e-logbook system makes data available to Scottish PO within two hours of the haul coming onboard. In other words, the e-logbook system in Scotland provides a substantial resource of real-time information that is relevant to quota species. What is missing from the current configuration of the e-logbook system in Scotland is about size structure of the catch and discarding. This type of information would be required if undersized fish (juveniles) were an important contribution to bycatch problem.

Sea State Inc also applies the principles of real-time reporting to demersal fisheries on the Pacific Northwest (off the coast of Washington and Oregon). The Pacific Northwest bycatch situation is very similar to the Scottish situation. In this region, shore-based trawlers have limited quota allocations of whiting and other demersal species. Discarding is not allowed at sea. As is the case in the North Sea there is a mismatch between fishing opportunity and restricted quotas that effectively creates choke species (e.g., rockfishes). Due to the unpredictable distribution of choke species in space and time they are difficult to avoid catching and the quota allocation for these species may be insufficient to cover what is caught by an individual vessel. The term “lightning strike” is used to refer to the possibility that one vessel will be unlucky and use up or exceed the entire bycatch allocation for the fleet in a single haul. A lightning strike haul has the potential to shut down the industry. In response to this risk, the fishing cooperatives have self-insured against “lightning strike” hauls by forming a risk pool. The relevance of this approach will be explored in greater detail in a FIS project (FIS 011B SMARTFISH: Selective management and retention of target fish).

5.5.2 Differences

The spatio-temporal distribution of salmon bycatch is semi-predictable in the sense that salmon undertake annual migrations and therefore the bycatch problem is narrowly constrained in both space and time. While they are migrating to natal rivers, salmon are captured by a high proportion of vessels (albeit to differing degrees as reflected by Tier 1,2,3 designation of cooperatives or vessels) rather than a lightning strike haul. Thus, the bycatch problem is likely to impact a high proportion of vessels. In the North Sea mixed fisheries a particular choke species may have a less predictable spatial and temporal distribution which is likely to be highly modified by fluctuations in abundance or climate-induced shifts in biogeography (e.g., hake). This might make the information less valuable particularly if there is not a high number of observations.

The EBS fisheries have the key advantage of 100% observer coverage of their fleet which is entirely funded by industry. The observer coverage in Scotland, which both MSS and the SFF conduct, is much less extensive. The MSS programme was originally designed (in 1978) from a stock assessment perspective, not a regulatory or enforcement one. MSS observers are not required to enforce a discard ban and it would greatly undermine their ability to collect scientific data should they be required to do so. The SFF, through SFF Services Limited, manages a team of fisheries observers who collect data for the Independent Onboard Observer Scheme which supports and informs the joint Industry/Government Fisheries Management and Conservation Group. The activities of this programme are growing and there is desire to utilise the resulting data in more ways than just informing the annual stock assessment process.

5.6 Conclusions

Reflecting on Sea State Inc.'s experience (> 25 years) with spatial management of EBS fisheries there would seem to be two distinct phases of developing improved information flow for reducing incidental bycatch in fisheries. Phase 1 is creating a mechanism for the data gathering, monitoring, reporting, and information sharing. This was the principle focus of part of the discussions with Sea State Inc (Section 3) Phase 2 is the more detailed statistical modelling, or post-processing of the bycatch data so as to generate predictive capability (Section 2). Phase 2 can both supplement the information generated in Phase 1. Furthermore, predictions from Phase 2 can be semi-continuously tested against data provided by Phase 1 creating a rolling validation that would reveal aspects of model performance.

Over the past two decades, the EBS fishing industry has passed through the developmental stage of Phase 1 but have not yet embarked on Phase 2. They have at their disposal a large, georeferenced database that can now be used for research purposes and to provide industry with some analytical insights into the dynamics of fish distribution over space and time. This should allow for a more profitable and efficient industry which offsets the costs of funding the data collection process (from observer programmes to contracting Sea State Inc). Access to world-leading fisheries scientists in the Seattle area could facilitate cutting-edge research into spatio-temporal dynamics of the EBS and Pacific Northwest fisheries.

In Scotland there is no impediment to Phase 1 and 2 developing in parallel, building on previous experience with real-time reporting and bycatch reduction in the EBS, the Pacific Northwest (Section 3), the east coast of the US (O'Keefe et al. 2014) and also the UK (Hetherington et al. 2015).

6 Consultation with the Scottish Stakeholders

During the course of the FISA project several formal and informal consultations with industry and Marine Scotland took place including:

- 24 June 2016 – meeting of Scottish Discard Steering Group (Aberdeen)
- 13 July 2016 – Industry and Policy Day, International Institute of Fisheries Economics and Trade (IIFET) Annual Conference (Aberdeen)
- 10 August 2016 – presentation given to the Scottish Industry Discards Initiative (Aberdeen)
- 23 August 2016 – break-out group discussion at the FIS Annual Scottish Fishing Conference (St. Andrews)
- 6 September 2016 – workshop to present the EBS experience to industry with Karl Haflinger and Eric Torgerson (Peterhead)
- 7 September 2016 - workshop to present the EBS experience to relevant stakeholders with Karl Haflinger and Eric Torgerson (Aberdeen)
- 8 September 2016 – meeting with David Anderson of Aberdeen Fish Producer's Organisation with Karl Haflinger and Eric Torgerson (Aberdeen)
- 30 September 2016 - meeting with Neil Campbell and Thomas Reilly of Marine Scotland - Science to discuss the observer and VMS databases
- 27 October 2016 – follow-up meeting with David Anderson of Aberdeen Fish Producer's Organisation
- October 2016 – follow-up discussions with several Shetland Producer's Organisations (conducted by Chevonne Angus of NAFC)

Detailed summaries from the two workshops (6th and 7th September 2016) will be written up separately for FIS011B so they will not be summarised here. Brief notes summarising the main points of the other meetings are given below.

6.1 Scottish Discard Steering Group

This day-long meeting was useful for informing the steering group about work being done by the University of Aberdeen and meeting representatives from a number of Scottish POs. Potential solutions for the landings obligation (LO) problem were identified as: gear selectivity, avoidance, flexibility and quota swapping. The latter was discounted as there is generally insufficient quota available for swapping for the species that are required, e.g. choke species. There are already some solutions being applied for gear as skippers change gear to avoid problems. Similar to the risk pool used in Pacific Northwest fishery (see Section 2.2.2 in FIS011B Report 2017) there are arrangements for banking and borrowing quota. Informal discussions with some of the skippers in attendance revealed that skippers are already sharing information about catch across a small network of peers via social media. This is evidence of the utility of the information for skippers. Despite this evidence that information about location of bycatch hotspots is desirable reservations were expressed about the principle of getting skippers to share information. This is somewhat contradictory viewpoint: real-time reporting is useful so we are doing it but it won't work more widely.

6.2 International Institute of Fisheries Economics and Trade

The issue of gathering and processing information was widely discussed at the IIFET conference during its Industry and Policy day. Several Scottish POs attended the Industry and Policy day, chairing and contributing to discussions. Advances in electronic reporting and monitoring were frequently mentioned in a number of presentations. The same ultimate aspiration was expressed by many of the international attendees. More timely information can transform the way that fisheries are managed. There are many opportunities to make greater use of current data as well as capture new types of data electronically. A significant challenge is to apply these data and technologies effectively in a policy and administrative sense. For example, collecting huge amounts of CCTV footage is becoming easier thanks to technology advances. A major challenge is efficiently processing this information and improving fisheries administration. Working through the processes and policies for managing fisheries data was felt to be as important as developing the technology to gather and process it.

International interest in real-time reporting were noted including the following:

- In New Zealand if the skipper can't cover catches with available quota by the year's end then a fine ("deemed value") is levied which is set at levels to encourage landing but discourage over-catch. Real-time information sharing is being explored as a means to reduce bycatch.

- In Pacific Northwest, information and communication technology is being used in demersal fisheries to do hotspot mapping very similar to that being used for Alaskan Pollock (Sylvia et al. 2014).
- Norway has had a discard ban in place since 1987 and uses a mix of move on regulations and RTCs. Fishers keep 20% of the landed value. Despite this, Norwegian discards remain significant.

A Cornwall-based skipper, David Stevens, made a key statement at the conference summing up his view of how to implement bycatch reduction: “Industry can’t push for flexibility without transparency” and “Science not enforcement”. Both statements are insightful and were referred to on several occasions.

A Swedish PO noted that, during efforts to reduce bycatch, the industry sought assistance of the Environmental Defence Fund. The industry felt strongly that fishermen needed to take responsibility for their own future. However, the “toolbox” that was required to help reduce bycatch was lacking.

6.3 Scottish Industry Discards Initiative

Prior to the Scottish Industry Discards Initiative (SIDI) meeting a briefing note describing the FISA project was distributed to attendees. The SIDI meeting had most of the major Scottish POs in attendance. A half-hour presentation summarising material in Section 3 was given in the afternoon (the morning had been taken up by the implications of Brexit). Post-Brexit in fact creates an opportunity of approaching the LO in a new way and customising a system for reducing bycatch that is based on Scottish experience. Questions were asked throughout the presentation and afterwards about how real-time reporting could work in practice in Scotland. Both positive and negative views were expressed about the potential for industry to share catch information about choke species. It was emphasised that the goal was not to share information about target species. Concern was expressed that under the LO the information could potential be used against them by Marine Scotland. The point was made that universities are currently incentivised to bring their expertise because under the Research Excellence Framework UK universities must demonstrate “impact” on UK industries. This creates an opportunity for partnerships such as those supported by FISA.

6.4 FIS Annual Scottish Fishing Conference

A breakout group discussion was held at the FIS Annual Scottish Fishing Conference to discuss the ongoing FIS project 011B (SMARTFISH: Selective management and retention of target fish). The group included representatives from industry (Tom Bryan-Brown of the Mallaig & North-West Fishermen's Association and one skipper Peter Bruce). The principle of real-time reporting was favourably perceived but the difficulties in convincing the industry that it was required were noted. The importance of skipper's personal knowledge was highlighted to underscore that technology only supplements the tactical decision making at sea. Concern was expressed that the information could be used against them by Marine Scotland. Another important consideration is that the Scottish industry is sensitive to implementing, even voluntarily, conservation measures which put them at a disadvantage relative to other nations fishing the same stocks. This illustrates how the information being shared by real-time reporting scheme needs to be viewed as valuable in its own right and not a burden on fishing.

6.5 Marine Scotland – Science

The meeting with Marine Scotland identifi ed the basic features of several geo-referenced fisheries databases in Scotland as well as their accessibility.

6.5.1 Logbook Information

The fishing logbook is the primary method of data collection. It records data on fishing operations by individual vessels by trip, and for each day of activity within a trip. These data are available since 1964 and include details of the catch, by species, in terms of the presentation and quantity of fish retained on board. Information is also collected on the fishing gear used and the area where the fish were caught. Area information are division, rectangle and zone as defined by the ICES. Council Regulations 1966/2006, 1006/2008 and 1224/2009 and Commission Regulations 1077/2008 and 201/2010, implemented by the Sea Fishing (EU Recording and Reporting Requirements) (Scotland) Order 2010 (SSI 2010/334), require Scottish vessels (when operating in Scottish, EU and third country waters) to record and report fishing activity data electronically. Software has been installed on board fishing vessels to record and submit data on fishing activities, with the expectation that electronic logbooks will eventually replace paper logbooks. Normally, catch data are submitted electronically within two hours of the haul coming on board.

All fishing activity data submitted electronically may be viewed by Marine Scotland Compliance. Primary vessel owners can also register on their systems, allowing them to view activity data for their vessels over an internet browser. The primary vessel owner can set up other users to view and administer their vessel activity data (POs or Agents). For scientific analysis of the data MSS aggregate landing data in space and time, remove any unique vessel or processor identifiers, and apply a disclosure limitation. Each aggregation must contain data from three or more vessels, and those aggregations with less than three vessels will be suppressed or, where appropriate, aggregated up to a higher spatial or temporal scale.

The electronic reporting systems operated by fishing vessels contain a limited range of validation checks to help ensure correct data are reported. In addition, to the validation processes, the information reported by fishermen is run through automatic cross-checks with other sources of information on activity available to fisheries administrations to ensure consistency and accuracy in the information reported. Landing declarations provide information on the weight and presentation of fish landed by species. Landing declarations and logbooks must be submitted to authorities within 48 hours of landing.

6.5.2 Observer Data

Digital data from scientific observers is available from mid-90s and are available upon request from Marine Scotland. These data do not have availability in real-time as the data are pre-processed and are available once the fishing trip finished. Observer data remove any unique vessel or processor identifiers. Between 60 and 90 trips are made annually to cover fishing operations in the North Sea and West of Scotland looking at whitefish or Nephrops (Table 1.1). Observers monitor the amount of each species caught and discarded, take measurements of the size composition and, for a selected group of species including cod, haddock and whiting, collect otoliths to determine the age of the fish. Following processing, these data are then submitted to ICES and combined with similar material collected in other countries to provide overall discard information.

6.5.3 VMS Data

VMS is a form of satellite tracking using transmitters on board fishing vessels. The system is a legal requirement under EC Regulation 2244/2003 and Scottish Statutory Instrument (SI) 392/2004. A basic VMS unit consists of a GPS receiver which plots the position of the vessel coupled with a communications device which reports the position at a minimum of every two hours. The unit automatically sends the following data on a pre-determined timescale: the vessel identification, geographical position, date/time of fixing of position, course and speed.

VMS data is considered personal data so access is strictly controlled. However, under the Data Protection Act vessel owners can request access to their VMS data in writing (by letter, fax or email). Vessel masters can also request VMS data for any period in which they can prove they were master of the vessel. From 2017 aggregate VMS data at the level of metiers will be available for analysis.

6.6 Aberdeen Fish Producer's Organisation

The Sea State Inc example highlights how real-time reporting has to be driven by industry needs. The fishing cooperatives (in the US) and POs (in Europe) have a central role to play. The FISA project therefore made a point of meeting different POs (e.g. through SIDI) and following up with more informal discussions with POs that seemed to be most interested in pursuing the idea. These included Aberdeen Fish Producer's Organisation (AFPO), Mallaig & North-West Fishermen's Association Ltd and several Shetland POs. Discussions with AFPO took place on three separate occasions during the course of the FISA project and were very helpful. A serious concern is the current inaccessibility of several of the relevant databases in Scotland (e.g., e-logbook, VMS and observer). A considerable amount of IT work would be required to improve accessibility of these data for uses described here. In the short term, funding support from Fisheries Innovation Scotland (FIS) could be sought to support this, in addition to any Marine Scotland funding. Over the short term this task could be supported by funding from the EMFF.

6.7 Shetland Fisheries

Following up on the workshops held in September, Dr Chevonne Angus (NAFC Marine Centre) met with several Shetland-based stakeholders including Shetland Fish Producer's Organisation (Brian Isbister) and Shetland Fishermen's Association (Simon Collins and Leslie Tait). In those discussions she summarised the EBS experience with a view to determining whether a similar means of real-time reporting could be developed locally. Industry representatives were interested in how the various data streams were used together for the real-time management. They could see a lot of potential, particularly if information from the seafood auction in Shetland could be integrated in some way. They acknowledged there was merit in learning from what is being done in the Bering Sea and cherry-picking and adapting parts of what is done there for local implementation. In general, fishermen are reluctant to share any data which may give away their 'hot spots' (NB this statement applies more to target species. Sharing information about choke species might be more acceptable). Concern was expressed that the information they shared could be used against them by Marine Scotland. This perception is partly based on their experience with RTCs. Brexit is the industry's immediate priority but in future it would be helpful to embark on discussions with skippers emphasising the potential for industry to design systems that work for them and that do not expose them to risk from a compliance perspective.

7 Recommendations

The modelling approach developed here was able to merge commercial fishing and survey data to predict spatio-temporal distribution of juvenile cod. One of the main strengths of this approach is the demonstration that it is possible to combine different sources of information about relative fish density comparable across time periods and areas by standardizing catch rates of different vessels and gears. This means that heterogeneity in catch rates caused by factors affecting catch rates are incorporated into the models via fixed or random effects, and that calibration between gears is estimated together with spatio-temporal changes in abundance. As a consequence, uncertainty in the calibration of a given gear against another is appropriately propagated in estimates of other parameter of the model. Two significant benefits of this approach are: 1) the ability to make effective use of many more data than is typically done by traditional single-metier/fleet analyses, and 2) where the contribution of a given metier to the data set is insufficient for cross-calibration to be reliably estimated, the corresponding data are automatically given less weight in the analysis.

7.1 Recommendations for Further Model Development

The modelling approach involves some simplifying assumptions which future developments of the model may want to explore more in depth. For example, fishing effort is assumed to be proportional (up to a constant) to the product of trawl width by trawling time. Secondly, total catch is assumed to be linearly related to fishing effort (essentially trawling time). Both assumptions are likely to be violated to some extent in reality. A further assumption of our models is that cumulated catch in one location does not affect subsequent local fish abundance. While it is clear that this assumption is untenable, it is unclear how problematic this is in practice, and unclear how it could be remedied until the majority of the catches is available to analyse.

Another important strength of the modelling approach is the use of high resolution Gaussian Markov Random Fields (GMRF), which allow the modelling of fine scale fish distribution hotspots. In addition, Bayesian inference permits a comprehensive incorporation of the uncertainty, because it takes into account error in the observations and in the latent field. However, the use of high resolution GMRF is computationally intensive, and thus, some trade-off had to be considered. We assumed that seasonal variation in distribution of juvenile cods is a large- rather than fine-scale spatial process (i.e. that seasonal changes affect large areas simultaneously). As a consequence, we were able to reduce the dimensionality (reflected in the number of parameters to be estimated) of the models by representing seasonal changes in spatial distribution with smoothing splines in order to reduce the computational burden associated with fitting the models. The splines were fitted as separate models without high-resolution spatial terms for increased speed. The fitted spline surface was then plugged-in as a fixed covariate in final models along with high-resolution spatio-temporal GMRF terms, ignoring estimation uncertainties about the seasonal components. Further developments of the model beyond the proof-of-concept stage should seek to properly account for such uncertainties, ideally by fitting the spline components as part of fitting the general model or alternatively by plugging in spline bases matrix as covariates in the model together with the multivariate-normal prior for the spline coefficients pre-computed from a previous analysis.

Despite the ability to combine data from several sources in our models, not all existing sources of data were used in the present work. We used data from the Scottish discard observation program, but a natural extension would be to use landings data, which are available at much finer temporal resolution, as all landing have to reported, while only a small subset of fishing trips have on-board observers.

A difficulty in using the combined VMS & Landings data lies in the uncertainty about where specific portions of the total landing have been caught along the recorded track of a vessel. Methods for assigning the catch spatially and probabilistically remain to be developed to make this source of data potentially informative at the spatial resolutions required by hotspot distribution models.

The next obvious, yet computationally challenging development of the approach will be the extension of our models to multiple species and multiple size classes, which will yet again increase the dimensionality of the models and therefore the computational complexity. Benefits should include 1) an expected substantial gain in power for the models especially in low fish density areas, thanks to borrowing information across several species and size classes; and 2) the ability to predict the expected costs and benefits for skippers to fish in a particular area across the whole community of relevant demersal species. These improvements would depend on having access to high performance computing facilities. Longer run times of more complex models also impact the ability of the predictions to be disseminated in real-time.

Finally, a more radical evolution would be to move from current phenomenological models to ones based on population dynamics processes (e.g. Kristensen et al. 2014). It is uncertain at this stage whether this would yield a net improvement or loss of predictive accuracy.

7.2 Recommendations for Real-Time Reporting

A review of the US EBS example (Section 3) indicates that the three basic components of a real-time reporting system are: (1) vessel-specific reporting about bycatch performance; (2) alerts indicating areas of high bycatch; and (3) maps showing locations of high, medium and low bycatch. In Scotland, the first component already exists. The e-logbook database would be capable of providing real-time information about the spatial location of high catches of quota species that are choke species, e.g. hake or saithe, in the situation where discarding was prohibited (e.g., a discard ban). Because the Scottish industry is required to upload catch information within two hours of the haul coming on board the turnaround time for observations to enter the database and become available for analysis is considerably better than the US. The e-logbook database reports spatial location of catches at the level of ICES statistical rectangle. This resolution is sufficient to proceed with developing the basic components of a reporting system: alerts and maps.

- Facilitate discussions with Marine Scotland and POs regarding on-going improvements to the functionality the e-logbook database with the aim of allowing POs to download and engage with their own data in a spatial sense.
- Identify an opportunity to develop a prototype reporting system that would include alerts about high bycatch of a choke species and possibly maps showing locations of high, medium and low bycatch.

- Explore the potential for VMS data being integrated in real-time with e-logbook database in ways that would satisfy industry and Marine Scotland requirements for confidentiality.

7.3 Recommendations for Institutional and Attitudinal Change

The recommendations in Section 5.2 depend on agreement across a range of stakeholders that real-time reporting is worth pursuing as a strategy for bycatch reduction. This agreement does not currently exist and there are serious obstacles which need to be overcome prior to getting consensus. The following recommendations address some of these obstacles:

- POs need to be motivated to change. The status of the landings obligation is in question and Brexit is imposing unprecedented uncertainty. On the one hand, this might limit the capacity of industry to consider other types of change (Section 4.7). Conversely, it could give the Scottish industry greater incentives to custom-build more effective means of bycatch reduction. This is consistent with the idea that industry needs to assume responsibility for solving the bycatch problem.
- More work needs to be done to convince industry of the advantages of real-time reporting. Information sharing is already happening informally on small scales using social media. Sharing across a trusted unit of collaboration, for example vessels belonging to a single PO, would increase the amount of information being shared thereby increasing the value of the information. A formal agreement, similar to the IPA used by the Pollock industry (see Section 3.2), would be helpful in defining information sharing protocols.
- Industry needs to be confident that implementing change (i.e., information sharing) will not expose them to risk of detection of non-compliance with fishing regulations. Marine Scotland needs to fully support industry adopting a co-management role. This is consistent with the greater emphasis being placed in results-based management of fisheries.
- As David Stephens (skipper of the Crystal Sea) noted: Industry can't push for flexibility without transparency. Flexibility needs to be accepted by Marine Scotland while transparency of reporting needs to be accepted by industry.

7.4 Enabling Funding

The recommendations above are very ambitious and would require substantial sources of funding. Several national sources exist, principally from FIS and FISA. The European Marine and Fisheries Fund is a source of funding to support activities that are supportive of the fishing industry that might be appropriate. Opportunistic sources (e.g., H2020) also exist. The statistical modelling described in Section 4 could potentially attract funding from research councils given the applicability of the model to industry. There is also potential for Knowledge Transfer Partnership funding for enabling research and co-development of tools at advantageous rates for the industry. It might also be useful to explore funding sources that can support further collaborations with US scientists working in Alaska and the Pacific Northwest (e.g., conservation-oriented foundations).

8 Acknowledgements

Funding from the Scottish Government through the Fishing Industry Science Alliance is gratefully acknowledged. Karl Haflinger and Eric Torgerson have been unfailingly generous in sharing their time and expertise. Chevonne Angus provided feedback about her meetings with POs in Shetland. Coby Needle suggested the initial work on the statistical modelling of juvenile cod which was undertaken by Max Lindmark, an MRes student in Applied Marine and Fisheries Ecology at the University of Aberdeen. David Anderson, Tom Bryan-Brown, Peter Bruce, Elena Balestri, Brian Isbister, Simon Collins and Leslie Tait have all been helpful in discussing the issues facing the Scottish industry.

9 References

Besag, J. 1974. Spatial interaction and the statistical analysis of lattice systems (with discussion). Journal of Royal Statistical Society, Series B, 36, 192–236.

Bailey, N., Campbell, N., Holmes, S., Needle, C. and Wright, P. 2010. Real-time closures of fisheries. Note for the Directorate General for Internal Policies, Policy Department B; Structural and Cohesion Polices, Fisheries. IP/B/PECH/IC/2009-091, 50 pp.

Bethoney, N.D., Schondelmeier, B.P., Stokesbury, K.D.E., and Hoffman, W.S. 2013. Developing a fine scale system to address River Herring (*Alosa pseudoharengus, A. aestivalis*) and American Shad (*A. sapidissima*) bycatch in the U.S. northwest Atlantic mid-water trawl fishery. Fisheries Research 141:79–87.

Bivand, R.S., Gómez-Rubio, V., and Rue,H. 2015. Spatial data analysis with R-INLA with some extensions. Journal of Statistical Software 63: 1–31.

Blangiardo, M., and Cameletti, M. 2015. Spatial and spatio-temporal Bayesian models with R-INLA. John Wiley & Sons, Ltd.

Boyd C., Woillez M., Bertrand S., Castillo R., Bertrand A., and Punt A.E. 2015. Bayesian posterior prediction of the patchy spatial distribution of small pelagic

fish in regions of suitable habitat. Canadian Journal of Fisheries and Aquatic Sciences. 72: 290-303.
Catchpole T.L., Frid C.L.J., and Gray T.S. 2005. Discards in North Sea fisheries: causes, consequences and solutions. Marine Policy, 29(5), 421-430.
Ciannelli, L., Fauchald, P., Chan, K.S., and Agostini, V.N., Dingsør, G.E. (2008). Spatial fisheries ecology: recent progress and future prospects. Journal of Marine Systems, 71(3), 223-236.
Gilks,W., Richardson, S., and Spiegelhalter, D. 1996. Markov Chain Monte Carlo in Practice. Chapman & Hall/CRC.
Consandey-Godin A., Krainski E.T., Worn B., and Flemming J.M. 2015. Applying Bayesian spatiotemporal models to fisheries bycatch in the Canadian Arctic. Canadian Journal of Fisheries and Aquatic Sciences. 72: 290-303.
Dunn D.C., Maxwell S.M., Boustany A.M., and Hapin P.N. 2016. Dynamic ocean management increases the efficiency and efficacy of fisheries management. Proceedings of the National Academy of Science. 113(3): 668-673.
European Commission. 1 January 2015: The landing obligation. (Page updated: February 17, 2015). Retrieved March 13, 2015, from http://ec.europa.eu/fisheries/cfp/fishing_rules/landing-obligation/index_en.htm
Gauvin, J.R., Haflinger, K., and Nerini. M. 1996. Implementation of a voluntary bycatch avoidance program in the flatfish fisheries of the eastern Bering Sea. University of Alaska, Alaska Sea Grant Program, Sea Grant Program Report 96-03, Fairbanks.
Gelman, A., Bois, F., and Jiang, J. 1996. Physiological pharmacokinetic analysis using population modelling and informative prior distributions. Journal of the American Statistical Association, 91, 1400–1412.
Hetherington, S.J., Bendall, V.A., de Rozarieux, N.A., Rodmell, D., and Stromberg, P. 2015. Stakeholder involvement in a UK led initiative to align spurdog management with the Landing Obligation under the reformed Common Fisheries Policy. Fisheries Science Partnership 2014-2015, Final report, 46 pp.
Holmes, S. J., Bailey, N., Campbell, N., Catarino, R., Barratt, K., Gibb, A., and Fernandes, P. G. 2011. Using fishery-dependent data to inform the development and operation of a co-management initiative to reduce cod mortality and cut discards. ICES Journal of Marine Science, 68: 1679–1688.
Ianelli, J.N., Honkalehto, T., Barbeaux, S., and Kotwicki S. 2015. Assessment of the walleye Pollock stock in the Eastern Bering Sea. http://www.afsc.noaa.gov/refm/stocks/plan_team/EBSPollock.pdf.
ICES. 2012. Manual for the International Bottom Trawl-surveys. Series of ICES Survey Protocols. SISP 1-IBTS VIII. 68 pp.

Jansen T., Kristensen K., Payne M., Edwards M., Schrum C., and Pitois S. 2012. Long-term retrospective analysis of mackerel spawning in the North Sea: A new time series and modeling approach to CPR data. PLOS ONE. 7(6) e38758.

Jaureguizar A.J., Wiff, R., and Luz C.M. Role of the preferred habitat availability for small shark (*Mustelus schmitti*) on the inter-annual variation of abundance in a large Southwest Atlantic Coastal System (El Rincón, 39-41 S). In press. Aquatic Living Resources.

Kristensen K., Thygesen U.H., Andersen K.H., and Beyer J.E. 2014. Estimating spatio-temporal dynamics of size-structured populations. Canadian Journal of Fisheries and Aquatic Sciences. 71: 326-336.

Little, A., Needle, C.L., Hilborn, R., Holland, D. and Marshall, C. T. 2014. Real-time spatial management approaches to reduce bycatch and discards: experiences from Europe and the United States. Fish and Fisheries. 16: 576-602.

Lichstein, J. W., Simons, T. R., Shriner, S. A., and Franzreb, K. E. 2002. Spatial autocorrelation and autoregressive models in ecology. Ecological monographs, 72(3), 445-463.

Madsen, S., and Haflinger, K. 2015. Chinook Salmon Bycatch Reduction Incentive Plan 2014. Report to National Marine Fisheries Service, April 2015. (https://alaskafisheries.noaa.gov/sites/default/files/reports/cpipa14.pdf).

Martino, S., and Rue, H. 2010. Case studies in Bayesian computation using INLA. *In* Complex data modelling and computationally intensive statistical methods. Pp 99-114

Needle, C. L. and Catarino, R. 2011. Evaluating the effect of real-time closures Oon cod targeting. ICES Journal of Marine Science 68: 1647-1655.

Nielsen J.R., Kristensen K., Lewy P., and Bastardie F.A 2014. A statistical model for estimation of fish density including correlation in size, space, time and between species from research survey data. PLOS ONE. 9(6) e99151.

Nielsen, K. N., Holm, P. and Aschan, M. 2015. Results based management in fisheries: Delegating responsibility to resource users. Marine Policy 51: 442-451.

Oritz, M. and Arocha, F. 2004. Alternative error distribution models for standardization of catch rates of non-target species from a pelagic longline fishery: billfish species in the Venezuelan tuna longline fishery. Fish. Res. 70(2-3): 275-297.

O'Keefe, C.E., DeCelles, G.R., Georgianna, D., Stokesbury, K.D.E. and Cadrin, S.X. 2010. Confronting the bycatch issue: an incentive-led approach to maximizing yield in the US Sea Scallop fishery. Pages 1–9 *in* Proceedings of the 2010 ICES Annual Science Conference. ICES, Copenhagen, Denmark.

O'Keefe, C., DeCelles, G.R., and Cadrin, S.X. 2014. Spatial and temporal patterns of Georges Bank yellowtail flounder from the SMAST bycatch avoidance program. Canadian Department of Fisheries and Ocean, Transboundary Resource Assessment Committee, Working Paper 2014/3.

Pelletier, D. and Ferraris, J. 2000. A Multivariate approach for defining fishing tactics from commercial catch and effort data. Can. J. Fish. Aquat. Sci. 57:51-65.

Rivoirard J., Simmonds J., Foote K.G., Fernandes P., and Bez N. 2000. Geostatistics for estimating fish abundance. Blackwell Science, Oxford, UK.

Roa-Ureta R., and Niklitschek E. 2007. Biomass estimation from surveys with likelihood-based geostatistics. ICES Journal of Marine Science. 64: 1723-1734.

Rue, H., Martino, S., and Chopin, N. 2009. Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion). Journal of the Royal Statistical Society, Series B, 71(2), 319–392.

San Martín, M., Wiff, R., Saavedra-Nievas, J.S., Cubillos, L.A. and Lillo, S. 2013. Relationship between Chilean hake (*Merluccius gayi gayi*) abundance and environmental variables in the central-southern of Chile. Fisheries Research. 143: 89-97.

Scottish Government. 2012. Scottish sea fisheries statistics 2011. A National Statistics Publication for Scotland, 99 pp. Scottish Government. (Page updated: Thursday, December 18, 2014). Retrieved March 13, 2015, from http://www.gov.scot/Topics/marine/Sea-Fisheries/discards.

Shelton A.O., Thorson J.T., Ward E.J., and Feist B.E. 2014. Spatial semiparametric models improve estimates of species abundance and distribution. Canadian Journal of Fisheries and Aquatic Sciences. 71: 1-12.

Stram, D.L., and Ianelli, J.N. 2015. Evaluating the efficacy of salmon bycatch measures using fishery-dependent data. ICES Journal of Marine Science 72: 1173-1180.

Sylvia, G., Cusack, C. and Swanson, J. 2014. Fishery cooperatives and the Pacific Whiting Conservation Cooperative: lessons and application to non-industrial fisheries in the western Pacific. Mar. Pol. 44: 65-71.

Thorson J.T., Scheuerell M.D., Shelton A.O., See K.E., Skaug H.J., and Kristensen K. 2015. Spatial factor analysis: a new tool for estimating joint species distributions and correlations in species range. Methods in Ecology and Evolution. 6: 627-637.

Ward, E.J., Jannot, J.E., Lee, Y.-W., Ono, K., Shelton, A.O. and Thorsen, J.T. 2015. Using spatiotemporal species distribution models to identify temporally evolving hotspots of species co-occurrence. Ecol. Appl. 25: 2198-2209.

Wiff R, Quiroz J.C, Tascheri R., and Contreras, F. 2008. Effect of fishing tactics on the standardization of cardinalfish (*Epigonus crassicaudus*) catch rates in the demersal multispecies fishery off central Chile. Ciencias Marinas. 34(2): 143-154.

10 Appendices

10.1 Appendix 1 R Code Developed

10.1.1 Appendix 1.1 Code for Pre-Processing the Data in Preparation for use in INLA

```
install.packages("INLA",
repos="http://www.math.ntnu.no/inla/R/stable")
install.packages("spdep")
install.packages("ggplot2")
install.packages("raster")
install.packages("rgdal")
install.packages("rgeos")
install.packages("RColorBrewer")
install.packages("alphahull")
install.packages("maptools")
source("http://peterhaschke.com/Code/multiplot.R")
###################
require(colorspace)
require(alphahull)
require(rgdal)
require(raster)
require(rgeos)
require(mgcv)
require(ggplot2)
require(grid)
require(sp)
require(spdep)
require(INLA)
require(RColorBrewer)
require(maptools)
###################

dev.off(dev.list()["RStudioGD"]) #to get rid of old plots
rm(list=ls()) #removing objects

################### loading data
ICES.shp <- shapefile("C:/Users/s04rw6/Aberdeen/ICES
areas/ICES_16th.shp")
UK.shp <- shapefile("C:/Users/s04rw6/Aberdeen/ICES
areas//map.shp")
data.fish <- read.csv("C:/Users/s04rw6/Aberdeen/msc
thesis//NS_IBTS.csv", header=T)
#data.obs.partial <-
read.csv("//home//rodrigo//Escritorio//Respaldo//PC//Aberdeen/
/msc thesis//Discard_clean.csv", header=T)
data.obs.rep <- read.csv("C:/Users/s04rw6/Aberdeen/ICES
areas/data_final_obs_rep.csv",header=T)
```

```
grid <- readAsciiGrid("C:/Users/s04rw6/Aberdeen/ICES
areas/bath_asc.txt")   # geographic depth
###################
data.obs.rep$Year=as.numeric(substring(data.obs.rep$Trip.ID,
2, 5))
data.obs.rep$month2=as.numeric(substring(data.obs.rep$Haul.Sta
rt.Date.and.Time, 4, 5))  #month from hauls when possible in
other case from landing date
ind.m=which(is.na(data.obs.rep$month2)=="TRUE")
data.obs.rep$month[ind.m]=data.obs.rep$Date.Landed...Month..nu
mber.[ind.m]
ID.rep=data.frame(ID=data.obs.rep$Trip.ID,Haul=data.obs.rep$Ha
ul.Number)
data.obs<-data.obs.rep[!duplicated(ID.rep),]
######### calculating average lat and long base on start/end
of trawl. When only one is available..keep it!
av.lon=NULL
av.lat=NULL
l=dim(data.obs)[1]
for(i in 1:l){
av.lon[i]=mean(c(data.obs$Haul.Dec.Start.Long[i],data.obs$Haul
.Dec.End.Long[i]),na.rm=T)
av.lat[i]=mean(c(data.obs$Haul.Dec.Start.Lat[i],data.obs$Haul.
Dec.End.Lat[i]),na.rm=T)
}
data.obs$ShootLon<-av.lon
data.obs$ShootLat<-av.lat
colnames(data.obs)[which(names(data.obs) == "Bottom.Depth")]<-
"depth2"
colnames(data.obs)[which(names(data.obs) ==
"FRS.Gear.Code")]<-"gear"
colnames(data.obs)[which(names(data.obs) ==
"Haul.Duration..min.")]<-"trawlmin"
colnames(data.obs)[which(names(data.obs) ==
"Haul.Duration..min.")]<-"trawlmin"
colnames(data.obs)[which(names(data.obs) == "Vessel.ID")]<-
"vessel"

################### getting rid of data with no coordinates
ind.na=which(data.obs$ShootLon=="NaN")
data.obs=data.obs[-ind.na,]
################### adding NA to depths <15 m
ind.z=which(data.obs$depth2<15)
data.obs$depth2[ind.z]<-NA
data.obs$source <- "observer"
data.fish$source <- "survey"
data.fish$trawlmin <- 30
data.fish$gear <- "MTD.sur"
colnames(data.fish)[which(names(data.fish) == "freq")] <-
"count"
```

```
m=max(data.obs$vessel,na.rm=T)  # giving the bigger number to
the survey vessel
data.fish$vessel <- m+1

#################### binding both databases
obs.subset <- subset(data.obs, select= c("ShootLon",
"ShootLat", "count", "trawlmin", "gear", "depth2", "month",
"Year", "source","vessel"))
fish.subset <- subset(data.fish, select= c("ShootLon",
"ShootLat", "count", "trawlmin", "gear", "depth2", "month",
"Year", "source","vessel"))
data.bind <- rbind(obs.subset, fish.subset)

data.bind$quarter=NA
############# creating "quaters"..spliting data into two
seasons
ind.m1=which(data.bind$mont<=6)
ind.m2=which(data.bind$mont>6)
data.bind$quarter[ind.m1]<-1
data.bind$quarter[ind.m2]<-2

############ cuting up the "historical" polygon
### leaving just north sea data
lon.min=-4.7
lon.max=7
lat.min=48
lat.max=61
data.bind.r <- subset(data.bind, ShootLon >= lon.min)
data.bind.r <- subset(data.bind.r, ShootLon <=lon.max)
data.bind.r <- subset(data.bind.r, ShootLat>=lat.min)
data.bind.r <- subset(data.bind.r, ShootLat<=lat.max)

##### getting rid of datapoint inland
ind.p<-which(data.bind.r$ShootLon < -1 & data.bind.r$ShootLat
< 54)
data.bind.r<-data.bind.r[-ind.p,]
ind.p<-which(data.bind.r$ShootLon < -4 & data.bind.r$ShootLat
< 58)
data.bind.r<-data.bind.r[-ind.p,]
ind.p<-which(data.bind.r$ShootLon < -2.1 &
data.bind.r$ShootLat < 55.8)
data.bind.r<-data.bind.r[-ind.p,]
data.bind.r<-data.bind.r[-2627,]

###################### cutting up just for Scotland 55 degrees
data.bind.r <- subset(data.bind.r, ShootLat >= 55)

#######################
###### generating perimeter of the data
Shoot.unique=unique(data.bind.r[c("ShootLon","ShootLat")])
Shoot.unique<-as.matrix(Shoot.unique)
```

```
alpha=0.4 #this control how "tight" is the perimeter around
the data
a.chull <- ahull(Shoot.unique,alpha=alpha)

per.lon=a.chull$arcs[,1]
per.lat=a.chull$arcs[,2]

####################################
#### to see perimiter around the datapoints
plot(per.lon,per.lat,type="l")
points(data.bind.r$ShootLon,data.bind.r$ShootLat,col="blue")
plot(UK.shp,add=T)
################# assign coors characters to perimetre data
crds <- cbind(x=per.lon,y=per.lat)
str(crds)
Pl <- Polygon(crds)
str(Pl)
ID <- "a.chull"
Pls <- Polygons(list(Pl), ID=ID)
str(Pls)
SPls <- SpatialPolygons(list(Pls))
str(SPls)
df <- data.frame(value=1, row.names=ID)
str(df)
perim <- SpatialPolygonsDataFrame(SPls, df)
str(perim)
################# extractig coordinates from ICES areas
ICES.poly.t<-ICES.shp@polygons
s.data=155488 #size of ICES database
med.lon=NULL
med.lat=NULL
for(i in 1:s.data){
  med.lon[i]=ICES.poly.t[[i]]@labpt[1]
  med.lat[i]=ICES.poly.t[[i]]@labpt[2]
}
med.points<-data.frame(lon=med.lon,lat=med.lat)

####### perimetre and datapoints of each cuadrucula with the
same coors structure
med.pt=data.frame(x=med.points$lon,y=med.points$lat)
coordinates(med.pt) = ~x+y
med.pt2=SpatialPoints(med.pt)
proj4string(med.pt2)<- CRS("+proj=longlat +datum=WGS84
+no_defs +ellps=WGS84 +towgs84=0,0,0")
proj4string(perim) <- proj4string(med.pt2)
######## Test which points fall within polygon
# takes arounds 10 min
win <- gWithin(med.pt2, perim, byid=TRUE)
ind.win=which(win=="TRUE")
med.points.cut=med.points[ind.win,]
```

```
######## adding coors structure
med.cut.coor=data.frame(x=med.points.cut$lon,y=med.points.cut$
lat)
coordinates(med.cut.coor) = ~x+y
med.cut.coor2=SpatialPoints(med.cut.coor)
proj4string(med.cut.coor2)<- CRS("+proj=longlat +datum=WGS84
+no_defs +ellps=WGS84 +towgs84=0,0,0")

######## cutting up cuadricula structure for my data
ind.cut<-over(med.cut.coor2,ICES.shp)
ICES.poly.win<-ICES.shp@polygons[unique(ind.cut$ET_ID)+1]
ICES.polygon.win<-SpatialPolygons(ICES.poly.win)
########### a plot to check
plot(ICES.polygon.win)
save.image(file = "C:/Users/s04rw6/Aberdeen/ICES
areas/workspace.RData")
```

10.1.2 Appendix 1.2 Code for Modelling Including GAM and INLA Model Fitting.

```
dev.off(dev.list()["RStudioGD"]) #to get rid of old plots
rm(list=ls()) #removing objects
###################
require(alphahull)
require(rgdal)
require(raster)
require(rgeos)
require(mgcv)
require(ggplot2)
require(grid)
require(sp)
require(spdep)
require(INLA)
require(RColorBrewer)
require(maptools)
require(reshape)
require(classInt)
##new laptop
load(file = "C:/respaldo/Aberdeen/ICES areas/m_year.RData")
load(file = "C:/respaldo/Aberdeen/ICES
areas/ICES.polygon.win2.RData")
load(file = "C:/respaldo/Aberdeen/ICES
areas/workspace_z.RData")
grid <- readAsciiGrid("C:/respaldo/Aberdeen/ICES
areas/bath_asc.txt")
################### specifying a subset of data in years (as
in msc thesis)
data.sub <- subset(data.bind.r, Year > 2010 & Year < 2016)
##################
data.sub<-data.sub[,-11]  ## getting rid of "quarter" colunm
```

```
################### tranforming datapoints in geographic
coordinates
pt=data.frame(x=data.sub$ShootLon,y=data.sub$ShootLat)
coordinates(pt) = ~x+y
pt2=SpatialPoints(pt)
proj4string(pt2)<- CRS("+proj=longlat +datum=WGS84 +no_defs
+ellps=WGS84 +towgs84=0,0,0")

#######

################## assigning datapoints to polygons in the
shape file
ind<-over(pt2,ICES.shp)
ID.po<-NULL
#d<-dim(ind.cut)[1]
d=2141  ### number of cuadricula in my cut area
for(i in 1:d){
  ID.po[i]=ICES.polygon.win2@polygons[[i]]@ID
}
ID.po=as.numeric(ID.po)

ind.points=match(ind$ET_ID,ID.po)
un=unique(ind.points)
ID.no.point=ID.po[-un]
ind.no.point=match(ID.no.point,ID.po)
########## making database with NULL observations for empty
cuadricula

data.sub$ICES.areas <- ind$ET_ID
data.sub$ICES.corr <- ind.points

################## making "year2" with correlative years
(allow to work with grouping data in INLA)
data.sub$Year2=NA
data.sub$YearMonth=NA
ye=sort(unique(data.sub$Year))
month.id=seq(1,12,1)
s=0
for(i in 1:length(ye)){
  ind=which(data.sub$Year==ye[i])
  data.sub$Year2[ind]=i
  for(j in 1:length(month.id)){
  ind2=which(data.sub$Year==ye[i] &
data.sub$month==month.id[j])
  data.sub$YearMonth[ind2]=j+s

  }
  s=12*i
}
################## making "vessels" a correlative number
data.sub$vessel2=NA
```

```
ve=sort(unique(data.sub$vessel))
for(i in 1:length(ve)){
  ind=which(data.sub$vessel==ve[i])
  data.sub$vessel2[ind]=i
}
###################### filling up each year-quarter with
missing quadricula
ID.po<-NULL
#d<-dim(ind.cut)[1]
ID2<-NULL
for(j in 1:d){
  ID.po[j]=ICES.polygon.win2@polygons[[j]]@ID

}
ID.po=as.numeric(ID.po)
month.id=seq(1,12,1)
ind.y=sort(unique(data.sub$Year2))
data.null<-c()
for(i in 1:length(ind.y)){
  for(j in 1:length(month.id)){
    ind.year.month=which(data.sub$Year2==ind.y[i] &
data.sub$month==month.id[j])

pt.m=data.frame(x=data.sub$ShootLon[ind.year.month],y=data.sub
$ShootLat[ind.year.month])
   coordinates(pt.m) = ~x+y
   pt2.m=SpatialPoints(pt.m)
   proj4string(pt2.m)<- CRS("+proj=longlat +datum=WGS84
+no_defs +ellps=WGS84 +towgs84=0,0,0")
   ind.m<-over(pt2.m,ICES.shp)
   ind.points.m=match(ind.m$ET_ID,ID.po)
   un.m=unique(ind.points.m)
   ID.no=ID.po[-un.m]
   ind.no.point=match(ID.no,ID.po)
   month.no=rep(month.id[j],length(ID.no))
   year.no=rep(ind.y[i],length(ID.no))

ID2=data.frame(ICES.areas=ID.no,Year2=year.no,month=month.no,I
CES.corr=ind.no.point)
   data.null=rbind(data.null,ID2)
    }
}
   s=0
data.null$YearMonth=NA
ye=sort(unique(data.null$Year2))
for(i in 1:length(ye)){
    ind=which(data.null$Year2==ye[i])
  for(j in 1:length(month.id)){
  ind.m=which(data.null$Year2==ye[i] &
data.null$month==month.id[j])
  data.null$YearMonth[ind.m]=j+s
```

```
  }
  s=12*i
}
##############  adding average depth in each ICES cuadricula
from raster
r <- raster(grid)
dep<- extract(r,ICES.polygon.win2)
ID.depth=NULL
for(i in 1:d){
  ID.depth[i]=as.numeric(ICES.polygon.win2@polygons[[i]]@ID)
}

data.sub$depth.geo=NA
di<-dim(data.sub)[1]
for(i in 1:di){
  pos=match(data.sub$ICES.areas[i],ID.depth)
  data.sub$depth.geo[i]=mean(subset(dep[[pos]],dep[[pos]]<0))
# if any value is positive...takes only negative in each ices
cuadricula
}

###########
dm=dim(data.null)[1]
data.null$depth.geo=NA

for(i in 1:length(ID.depth)){
   ind.nul=which(data.null$ICES.areas==ID.depth[i])

data.null$depth.geo[ind.nul]=mean(subset(dep[[i]],dep[[i]]<0))
}

le.nul=dim(data.null)[1]
lon.nul<-NULL
lat.nul<-NULL
for(j in 1:le.nul){
id<-data.null$ICES.corr[j]
lon.nul[j]=ICES.polygon.win2@polygons[[id]]@labpt[1]
lat.nul[j]=ICES.polygon.win2@polygons[[id]]@labpt[2]
}
data.null$ShootLon=lon.nul
data.null$ShootLat=lat.nul
data.null$count=rep(NA,le.nul)
data.null$trawlmin=rep(120,le.nul)
data.null$gear=rep(NA,le.nul)
data.null$Year=rep(NA,le.nul)
data.null$source=rep("nule",le.nul)
data.null$vessel=rep(NA,le.nul)
data.null$vessel2=rep(NA,le.nul)
data.null$depth2=rep(NA,le.nul)
#############
####### binding positive and NULL observations
```

```
data.sub.bin=rbind(data.sub,data.null)

############ binning depth.geo data in each 10 m (delta) for
performing rw2
min.dg=min(data.sub.bin$depth.geo,na.rm=T)-0.1  #to be sure to
include the boundaries
max.dg=max(data.sub.bin$depth.geo,na.rm=T)+0.1
delta=10
sq=seq(min.dg,max.dg,delta)
data.sub.bin$depth.geo.bin=data.sub.bin$depth.geo
for(i in 1:(length(sq)-1)){

ind.dg=which(data.sub.bin$depth.geo>=sq[i]&data.sub.bin$depth.
geo<sq[i+1])
  data.sub.bin$depth.geo.bin[ind.dg]=round((sq[i]+sq[i+1])/2)
}
############ also adding a max lim for fishing of 200 m,
############ acording to the observe depth in loogbooks
(depth2)
ind.dep=which(data.sub.bin$depth.geo.bin< -200)
data.sub.bin$depth.geo[ind.dep]=-200
#########################
data.sub.bin$depth.geo.bin=round(data.sub.bin$depth.geo.bin) #
help precision
################### PREGAM
### jusy apply pregam to observed data
ind.data<-which(data.sub.bin$source=="observer" |
data.sub.bin$source=="survey")
data.pregam<-data.sub.bin[ind.data,]

#### to run the historical pregram

pregam<-gam(count ~
ti(month,depth.geo.bin,bs=c("cc","tp"))+te(ShootLon,ShootLat,m
onth,bs="ts")+te(ShootLon,ShootLat,by=month,bs="re")+te(ShootL
on,ShootLat,bs="re")+ti(depth.geo.bin,bs="tp")+ti(month,bs="cc
")+gear+Year2+offset(log(trawlmin)), data= data.pregam,
family= poisson)
################ marginal for month/depth
#marginal depth and space
marg.data <- data.frame(depth.geo.bin=
data.pregam$depth.geo.bin,

ShootLon=data.pregam$ShootLon,

ShootLat=data.pregam$ShootLat,
                                    month=data.pregam$month,
                                    Year2=3,
                                    gear="MTD.sur",
                                    trawlmin=120)
```

```
marg<- predict(pregam, newdata = marg.data, type = "response",
se = TRUE)
depth.month.marg.data <- data.frame(depth.geo.bin=
data.pregam$depth.geo.bin,

ShootLon=mean(data.pregam$ShootLon,na.rm=TRUE),

ShootLat=mean(data.pregam$ShootLat,na.rm=TRUE),
                              month=data.pregam$month,
                              gear="MTD.sur",
                              Year2=3,
                              trawlmin=120)

depth.marg<- predict(pregam, newdata = depth.month.marg.data,
type = "response", se = TRUE)
space.marg.data <- data.frame(depth.geo.bin=
mean(data.pregam$depth.geo.bin,na.rm=TRUE),

ShootLon=data.pregam$ShootLon,

ShootLat=data.pregam$ShootLat,
                                     month=data.pregam$month,
                                     gear="MTD.sur",
                                     Year2=3,
                                     trawlmin=120)

space.marg<- predict(pregam, newdata = space.marg.data, type =
"response", se = TRUE)
Year.marg.data <- data.frame(depth.geo.bin=
mean(data.pregam$depth.geo.bin,na.rm=TRUE),

ShootLon=mean(data.pregam$ShootLon,na.rm=TRUE),

ShootLat=mean(data.pregam$ShootLat,na.rm=TRUE),
                                     month=6,
                                     gear="MTD.sur",
                                     Year2=data.pregam$Year2,
                                     trawlmin=120)

Year.marg<- predict(pregam, newdata = Year.marg.data, type =
"response", se = TRUE)
data.sub.bin$pregam.depth<-depth.marg
data.sub.bin$pregam.space<-space.marg
data.sub.bin$pregam.depth<-NA
data.sub.bin$pregam.space<-NA
data.sub.bin$pregam.Year<-NA
data.sub.bin$pregam.depth[ind.data]<-depth.marg$fit
data.sub.bin$pregam.space[ind.data]<-space.marg$fit
data.sub.bin$pregam.Year[ind.data]<-Year.marg$fit
###################################  end of pregam
###### Constructig a database for prediction in each Year
```

```
ind.y=sort(unique(data.sub.bin$Year2))
data.pred<-c()
un=unique(data.sub.bin$ICES.corr)
le.pred=length(un)
lon.pred<-NULL
lat.pred<-NULL
for(g in 1:length(un)){
  id<-un[g]
  lon.pred[g]=ICES.polygon.win2@polygons[[id]]@labpt[1]
  lat.pred[g]=ICES.polygon.win2@polygons[[id]]@labpt[2]
}


s=0
for(i in 1:length(ind.y)){
  for(j in 1:length(month.id)){
  pred.ICES.corr=rep(un,1)
  pred.source=rep("prediction",le.pred)
  pred.depth2=rep(NA,le.pred)
  pred.Year=rep(NA,le.pred)
  pred.trawlmin=rep(120,le.pred)  # predictions per two hour
of trawling
  pred.count=rep(NA,le.pred)
  pred.depth.geo=NA
  pred.ICES.areas=NA
for(f in 1:length(un)){
  ind.a=which(data.sub.bin$ICES.corr==un[f])
  pred.ICES.areas[f]=data.sub.bin$ICES.areas[ind.a[1]]
  ind.d=match(pred.ICES.areas[f],ID.depth)
  pred.depth.geo[f]=mean(subset(dep[[ind.d]],dep[[ind.d]]<0))
  }
pred.gear=rep(NA,le.pred)
pred.month=rep(j,le.pred)
pred.Year2=rep(i,le.pred)
pred.ShootLon=lon.pred
pred.ShootLat=lat.pred
pred.vessel=rep(NA,le.pred)
pred.vessel2=rep(NA,le.pred)
pred.pregam.depth=rep(NA,le.pred)
pred.pregam.space=rep(NA,le.pred)
pred.pregam.Year=rep(NA,le.pred)
pred.YearMonth=rep(j+s,le.pred)
data.pred_it=data.frame(ShootLon=pred.ShootLon,
ShootLat=pred.ShootLat, count=pred.count,
trawlmin=pred.trawlmin, gear=pred.gear, depth2=pred.depth2,
month=pred.month, Year=pred.Year,Year2=pred.Year2,
source=pred.source,ICES.areas=pred.ICES.areas,ICES.corr=pred.I
CES.corr,vessel=pred.vessel,YearMonth=pred.YearMonth,
depth.geo=pred.depth.geo,
vessel2=pred.vessel2,pregam.space=pred.pregam.space,pregam.dep
th=pred.pregam.depth,pregam.Year=pred.pregam.Year)
```

```
data.pred=rbind(data.pred,data.pred_it)
  }
s=12*i
}
data.pred$depth.geo.bin<-NA
####### binding positive, NULL observations, and predictions
data.sub.bin.p=rbind(data.sub.bin,data.pred)
data.sub.bin.pred<-data.sub.bin.p
############ binning depth.geo data in each 10 m (delta) for
performing rw2
min.dg=min(data.sub.bin.pred$depth.geo,na.rm=T)-0.1  #to be
sure to include the boundaries
max.dg=max(data.sub.bin.pred$depth.geo,na.rm=T)+0.1
delta=10
sq=seq(min.dg,max.dg,delta)
data.sub.bin.pred$depth.geo.bin=data.sub.bin.pred$depth.geo
for(i in 1:(length(sq)-1)){

ind.dg=which(data.sub.bin.pred$depth.geo>=sq[i]&data.sub.bin.p
red$depth.geo<sq[i+1])

data.sub.bin.pred$depth.geo.bin[ind.dg]=round((sq[i]+sq[i+1])/
2)
}
############ also adding a max lim for fishing of 200 m,
############ acording to the observe depth in loogbooks
(depth2)
ind.dep=which(data.sub.bin.pred$depth.geo.bin< -200)
data.sub.bin.pred$depth.geo.bin[ind.dep]=-200
#########################
data.sub.bin.pred$depth.geo.bin=round(data.sub.bin.pred$depth.
geo.bin) # help precision

################### selecting and cutting up layers for
prediction
################### getting rid of all layers first
ind.pred=which(data.sub.bin.pred$source=="prediction")

data.pred2<-data.sub.bin.pred[ind.pred,]
data.sub.bin.pred=data.sub.bin.pred[-ind.pred,]
##################
################## selecting only month 2 and 8 (more survey
data)#### and a year in the middle
ind.cut2<-which(data.pred2$month==2 | data.pred2$month==8)
data.pred3<-data.pred2[ind.cut2,]
ind.cut3<-which(data.pred2$Year2==4)   ###### taking 2014 for
year round predictions
data.pred4<-data.pred2[ind.cut3,]   #### when I also need a
layer for 1 extra year
data.pred2<-rbind(data.pred3,data.pred4)
```

```
############### cutting up the area

lon.max=3
lat.min=57
lat.max=60

data.pred.cut <- subset(data.pred2, ShootLon <= lon.max)
data.pred.cut <- subset(data.pred.cut, ShootLat>=lat.min)
data.pred.cut <- subset(data.pred.cut, ShootLat<=lat.max)
data.pred.cut2 <- subset(data.pred2, ShootLon <= lon.max)
data.pred.cut2 <- subset(data.pred.cut2, ShootLat>=lat.min)
data.pred.cut2 <- subset(data.pred.cut2, ShootLat<=lat.max)
######################
##################
data.sub.bin.pred2=rbind(data.sub.bin.pred,data.pred.cut2)  #
only months predictic layers

ICES.nb <- poly2nb(ICES.polygon.win2) #neighbours structure

################## plotting nb structure
nb2INLA("LDN.graph", ICES.nb)
LDN.adj <- paste(getwd(),"/LDN.graph",sep="")
H <- inla.read.graph(filename="LDN.graph")
image(inla.graph2matrix(H),xlab="",ylab="")

################# some definitions
data.sub.bin.pred2$ICES.corr1<-data.sub.bin.pred2$ICES.corr #a
repetition needed for INLA
data.sub.bin.pred2$ICES.corr2<-data.sub.bin.pred2$ICES.corr #a
repetition needed for INLA
data.sub.bin.pred2$ICES.corr3<-data.sub.bin.pred2$ICES.corr #a
repetition needed for INLA
data.sub.bin.pred2$quarter<-NA
ind.q1=which(data.sub.bin.pred2$month<=6)
ind.q2=which(data.sub.bin.pred2$month>=7)
data.sub.bin.pred2$quarter[ind.q1]=1
data.sub.bin.pred2$quarter[ind.q2]=2
data.sub.bin.pred2$quarter2<-NA
ind.q1=which(data.sub.bin.pred2$month<=3)
ind.q2=which(data.sub.bin.pred2$month>=4
&data.sub.bin.pred2$month<=6)
ind.q3=which(data.sub.bin.pred2$month>=7
&data.sub.bin.pred2$month<=9)
ind.q4=which(data.sub.bin.pred2$month>=10)
data.sub.bin.pred2$quarter2[ind.q1]=1
data.sub.bin.pred2$quarter2[ind.q2]=2
data.sub.bin.pred2$quarter2[ind.q3]=3
data.sub.bin.pred2$quarter2[ind.q4]=4
################## selecting the block
```

```
ind.block1<-which(data.sub.bin.pred2$YearMonth>36 &
data.sub.bin.pred2$YearMonth<=48)
data.sub.bin.pred.block1<-data.sub.bin.pred2[ind.block1,]
data.sub.bin.pred.block1$count=round(data.sub.bin.pred.block1$
count)
data.sub.bin.pred.block.p<-data.sub.bin.pred.block1
save(data.sub.bin.pred.block.p, file =
"C:/respaldo/Aberdeen/ICES
areas/data.sub.bin.pred.block.p.RData")

################  priors from block before


########### fix effects
name=model.1.res  # define the name of the model to use as
prior
fixed=name$summary.fixed
fixed.mean=fixed$mean[2:dim(fixed)[1]] # taken all fix but
intercept
fixed.sd=fixed$sd[2:dim(fixed)[1]]
fixed.prec=1/(fixed.sd)^2 # acording to the INLA book
definition of precision (tau).
fixed.names=name$names.fixed[2:dim(fixed)[1]]
f<-paste(fixed.names,"=",fixed.mean,collapse=" , ")
f2<-paste(fixed.names,"=",fixed.prec,collapse=" , ")
f.mean<- eval(parse(text=paste('list(', f, ')')))
f.prec<- eval(parse(text=paste('list(', f2, ')')))
control.f=list(mean=f.mean,prec=f.prec,expand.factor.strategy
= "inla")

############ priors for random effect
############
rnd=name$summary.hyperpar
rnd.mean=rnd$mean
rnd.sd=rnd$sd
rnd.prec=1/(rnd.sd)^2
rownames(rnd)
# [1] hyper for besag, [2] hyper for autocorr, [3] hyper for
iid
hyper.besag <-list(prec=list(prior="loggamma",
params=c(rnd.mean[1], rnd.prec[1])))
hyper.autocorr <-c(rnd.mean[2],rnd.prec[2])
hyper.iid<-list(prec=list(prior="loggamma", c(rnd.mean[3],
rnd.prec[3])))

model.1 <- count ~ pregam.depth+ pregam.space+ f(ICES.corr,
model = "besag", graph = LDN.adj, hyper=hyper.besag,
group=YearMonth, constr=T,
control.group=list(model="ar1",hyper=hyper.autocorr))+month+Ye
ar2+gear+f(vessel2,hyper=hyper.iid)+offset(log(trawlmin))
```