**Abstract**

**Introduction**

Significant and enduring challenges have complicated efforts to identify and understand the ecological effects of the Exxon Valdez oil spill on marine communities. These include i) a general lack of pre-spill data that restricts direct before-after comparisons of specific species or communities and ii) a substantial amount of spatial and habitat-driven variation of species and communities that confound direct comparison of oil exposed and unexposed areas.

It is well known that exposure to petroleum contamination can have wide-ranging effects on marine species. Documented effects of petroleum range from the acute and direct (e.g. toxicological effects that cause serious injury or immediate death) to the gradual and indirect (modifications to physiological pathways and reproductive characteristics). In recent years, a consensus has developed indicating that even relatively low levels of exposure can have considerable fitness consequences for individuals (Collier et al. 1993, hicken, incardona). While detailed laboratory studies have repeatedly found negative consequences of petroleum related contaminants, connecting these excellent individual level studies to population or community level effect in natural populations has proven exceedingly difficult (New Eco Apps paper Awkerman et al 2016).

The Exxon Valdez Oil Spill (hereafter “EVOS”) in March 1989 spilled approximately 257 thousand barrels (36,000 mt) of crude oil into Prince William Sound, Alaska. Over the following weeks and months the oil was advected by wind and currents to the southwest, spreading into the Gulf of Alaska. While the exact dispersal path and final distribution of oil in water and sediments remains uncertain, the approximate distribution and extent of surface oil is documented (Wolfe et al). Multiple lines of evidence suggest that oil from EVOS reached a wide range of habitats – oil was observed both directly in some deep water trawls (Armstrong et al. 1995) and a range of elevated petroleum metabolites were detected in a wide range of invertebrate and fish species from a wide spatial area (Collier et al. 1993, Anderson et al. 1995, Krahn et al. 1992, Marty et al. 2003, Sol et al. 2000). In at least some habitats EVOS oil persists in the environment: oil has been documented nearly 30 years after the spill along shorelines (Short et al. 2007, Peterson et al 2003).

Despite undisputed evidence that oil spills have negative consequences for marine ecosystems () and that EVOS was a widely dispersed and disruptive event in the Gulf of Alaska (Peterson et al. 2003), it has been difficult to quantify ecosystem wide consequences of EVOS. Certainly direct mortality to marine mammals (Garrott et al. 1993, Doroff et al. 1993), birds (Piatt et al. 1990, Piatt and Ford 1996), and other highly visible and charismatic species that were directly exposed to crude oil is well documented. However, ecological systems are extraordinarily complex systems and theory suggests that many ecosystem level consequences of an event like EVOS are the result of indirect interactions and complex species-by-environment interactions (Wootton refs, others). Thus, a retrospective examination of the ecological consequences of EVOS 35 years later is warranted.

Here we focus on demersal fish communities of the central and western Gulf of Alaska and use available information to explore community responses to the EVOS spill. Demersal fish communities were exposed to EVOS as oil swept west out of Prince William Sound (Fig. 1, Collier et al. 1993, Sol et al. 2000). While the extent and exposure of fish communities to oil is almost wholly unknown, it is clear that some areas were exposed to oiling while other, similar habitats were left unexposed. We leverage this spatial gradient in exposure to contrast demersal fish communities over the past 40 years (1984-2015). Specifically, we develop and apply a suite of spatio-temporal models to the groundfish survey data, assess multiple metrics of demersal fish communities, and compare areas that experienced a range of exposure to EVOS to identify any potential EVOS related signal in changes to the groundfish community.

**Methods**

*Data*

To assess spatial and temporal patterns in groundfish communities we used fisheries independent surveys conducted by the Alaska Fisheries Science Center (AFSC) between 1984 and 2015. This AFSC dataset has used the same methodology over the time series (stratified random sampling design, with the same bottom trawl sampling gear throughout the time series; Stauffer (2004) andArmistead (1993). The average latitude and longitude of survey tows were used to represent the spatial locations of each sample, and these values were converted to ALBERS PROJECTION which is appropriate for the Gulf of Alaska (need to dig out methods for Albers projection). As our interest was primarily in the waters most likely affected by EVOS, we restricted our analysis to trawls conducted shallower than 600m. In total, we included 7601 individual hauls between 1984 and 2015.

We applied our models to 55 species groups. These groups represent species that were observed in at least 270 survey tows. 53? represent individual species and YY represent species complexes that were not identified to species separately during field sampling throughout the survey extent (supplementary TABLE). For brevity, we refer to these species and species groups simply as “species” subsequently. We provide the full species list and number of tows in which each species was observed in the supplement.

Things to do for Data:

Make map of trawl locations for each year.

· Write detailed methods for how I separated out the different sizes

o Dealing with no observed size data.

o Converting observed length to weight using allometric equations.

o Determining which fraction are “big” versus “small” and applying that to the CPUE data to generate CPUE for each size class.

*Statistical modeling*

We constructed separate models for each groundfish species to understand the spatiotemporal patterns of occurrence and abundance. We estimated a model for each species independently and subsequently combined the model outputs to generate a suite of multi-species metrics of the groundfish communities. We first present the spatio-temporal statistical model and then describe the spatio-temporal metrics of community change.

In all models, we used catch per unit effort (kg∙hectare-1) observed on each trawl as the response variable. Because most species were absent from a large number of observed trawls, we separately modeled the presence and distribution of species density, adopting a delta-GLM approach with two sub-models (Pennington 1983, Maunder and Punt 2004, Shelton et al. 2014). Probability of occurrence of the *i*th species in year *t* for a set of locations ***s***, ***π****it*(***s***), was modeled using a binomial GLMM with logit link,

*logit*[***π****it*(***s***)] = ***X****t*(***s***) ***β***i+*εit*(***s***) (1)

where ***X****t* (***s***) represents a matrix of fixed effect covariates at locations ***s*** in year *t*, ***β***i represents a vector of coefficients for species *i*, and ***ε****it*(***s***) represents a vector of spatial random effects that follow a first-order autoregressive process:

***ε****it*(***s***)~*Normal*[***ρεi****t-1*(***s***),***Σ****i*(***s)***] (2)

where ***ρ*** represents the degree of autoregression in encounter probabilities and ***Σ*** represents spatial covariation in random effects (discussed below). Random effects were assumed to be autoregressive to account for variation not explicitly included in our model (e.g., variation due to the environment or population processes such as density dependence). Note that because trawl surveys are not conducted annually but triennially (1984-1999) or biennially (1999-2015) the autoregressive term refers to the date of the previous survey year, not the previous calendar year.

For the second groundfish sub-model (hereafter positive model), we assumed that groundfish CPUE was gamma distributed and used a log-link,

*Yit* (***s***)~*Gamma*(***μ***it (***s***),*σi*)  
 *log*[***μ****it* (***s***) ]= ***Z****t*(***s***) ***γ****i*+***η****it* (***s***) (3)

where ***Z****t*(***s***) is a matrix of covariates corresponding to each haul location, ***γ****i* represents the estimated species-specific coefficients, ***η****it* (***s***) represent spatial random effects that again follow an autoregressive process (analogous to eqn 2, but with a independently derived covariance matrix). Then ***μ****it* (***s***) represents the mean of the gamma distribution and *σi* represents the **shape** parameter of the gamma distribution. This model only uses the hauls in which the species was observed and so the positive model describes the CPUE of each species conditioned on the species presence.

For all models we used available haul level information about bottom depth (m) and included both linear and quadratic terms for log(depth) in the occurrence and positive models (Shelton et al. 2014, Tolimieri et al. 2015). We centered log(depth) by subtracting the mean. We also considered two models for fixed year effects: in one model we estimated a single intercept while in the other we allowed for a distinct intercept for each year. As the intercept scales the occurrence (or CPUE, respectively) for the entire region, models with variable intercepts allow for spatially uniform, region wide changes in occurrence or CPUE.

For the spatial random effects in the occurrence and positive models, we used the Matern function to model covariance as a function of Euclidian distance, so ***Σ***(*sj*,*sk*) = *τ*2/*Γ*(*ν*)2ν-1 (*κdjk*)ν *K*ν(*κdjk*), is the covariance between location *sj* and *sk*, *τ*2 is the spatial variance, *Γ*() and *K*ν() represent the Gamma and Bessel functions, respectively, *djk* is the Euclidian distance between locations *sj* and *sk*, and *κ* is an estimated scaling parameter (Lindgren et al. 2011). The parameter *ν* controls the smoothness of the Matern function and is usually fixed rather than estimated from data (when *ν*  = 0.5, the Matern reduces to the exponential covariance function). Following previous work, we chose *ν* = 3/2; this allows the Matern to be more flexible than the exponential, but also allows the function to be differentiable (Rasmussen & Williams 2006, Ward et al 2015). The covariance matrices for the presence-absence and positive models for each species have separate parameters *τ*2 and *κ*, reflecting the assumptions that each model component may have a different variance or rate at which correlations decline as a function of distance. Further details can be found in Ono et al. (2015) and Ward et al. (2015).

We estimate the model using the integrated nested Laplace approximation as implemented in the R package INLA (www.r-inla.org, Martins et al. 2013). INLA approximates the inverse of the spatial variance-covariance matrix of fixed locations using three large sparse matrices using stochastic partial differential equations (Rue et al. 2009, Ruiz-Cardenas et al. 2012). Estimation of the fixed effects is then done via marginal maximum likelihood using the Laplace approximation to approximate the integral across random effects, and random effects are estimated via Empirical Bayes. Using these estimates (and Bayesian priors on fixed effects), INLA allows Monte Carlo samples to be generated from the posterior distribution, as the Laplace approximation to the marginal likelihood.

For each species, we estimated two occurrence models and two positive models. For each submodel, we estimate one model with a single intercept and one with a year-specific intercept. We compare single and year-specific models using posterior predictive plots and deviance information criterion (DIC) and then identified preferred models for each species.

*Generating predictive densities for each species*

After the two sub-models were estimated for each species, we used the estimated models to generate predicted densities for Gulf of Alaska. We projected our model estimates to the center of 2x2 km grid created for the entire Gulf of Alaska (add details of the projection to a supplement). We generated 1,000 Markov Chain Monte Carlo (MCMC) samples from the joint approximate posterior density for each species and for each MCMC sample we predicted a density for each sub-model to the 2x2km grid. We then combined the occurrence and positive models to generate an unconditional expectation for CPUE for each grid cell. Using MCMC samples from the full posterior distribution maintains the spatio-temporal correlation structure of the estimated parameters and random effects and properly accounts for uncertainty in these estimates. Because the occurrence and positive models are estimated independently, we can calculate the unconditional expectation for CPUE of species at time *t* and location *s* by multiplying each MCMC sample from the occurrence and positive model. Specifically, for the *g*th MCMC sample, the unconditional CPUE estimate is *πit*g(*s*) ∙ *μ*itg (*s*) in units kg∙hectare-1.

*Defining areas for comparison across the Gulf of Alaska*

We identified eleven areas across the Gulf of Alaska to compare groundfish communities through time (Fig. 1). Each area represents habitat between 50 and 150m deep divided by natural bathymetric breaks (canyons). Areas ranged from X to Y km2 (Table YYY). We ensured that each area was separated by at least THIS MANY km.

*Community metrics*

For each area, we summarized the groundfish community by constructing 10 community metrics. We summarize the predicted CPUE for each species in each year in each region using the MCMC draws for each 2x2 km grid cell. We then combine information across grid cells within each area to produce a mean estimate (and uncertainty) for CPUE within each area. Thus for each area, we have a time-series of density for each species for survey years between 1984 to 2015. We then combine these species-specific metrics to generate multi-species community metrics for each area in each year (see below). These time-series allow for

Total biomass. Simple sum of all species in the sample.

Guild biomass:

Apex predators (piscivore), Benthivore, pelagic

Feeding Groups:

Taxonomic Groups:

Sharks vs. non-Sharks.

Diversity.

We assessed three biodiversity indices: species richness, Shannon diversity (H’), and Simpson’s index of species dominance (D). We calculated species richness for each area as the summed occurrence of all species in an area. We calculated H’as the sum of proportional biomass (CPUE) of all species in an area:, where is the proportional abundance of species in an area. The Simpson’s index was quantified as . We converted H’ and D to effective species diversity (or true diversity) by exp(H’) and 1/D, respectively (see Jost papers for rationale in using effective measures. Also need to decide on whether to use occurrence data or biomass data - maybe do both and include some results in supplementary materials). These latter two diversity metrics were estimated in the R package vegan (Oksanen 2016). In addition, we used Rao’s quadratic index (Rao 1982) to quantify the functional and taxonomic diversity of demersal species in each area, using life history and trophic characteristics of each species or similar species (supplementary table). Rao’s Q was calculated in the R package XX (ref).

Others

Calculate these metrics:

1. Linear Trend (weighted by inverse of variance)

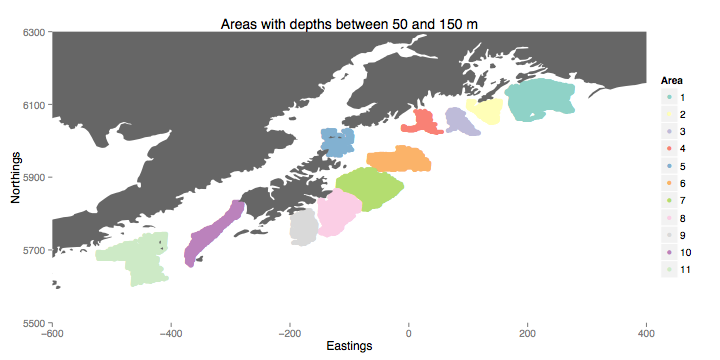
2. Grand Mean Density (weighted by inverse of variance)

3. Temporal Variance around trend in each area

4. Total variance across area through time

Plots:

Plot time series for each area (highlight areas expected to be most influenced by oil), look for change around 1989.



Each of the numbered circles is a shallow area that should have approximately similar fish communities based on physical parameters (50 to 150m) and they were all surveys before the exxon spill (1984 and 1987) and after the spill (1990 and on). They experienced different levels of direct exposure to the oil spill, though. We would expect that oil spill effects would be most intense near the oil (in space and time) and decline as a function of both distance from the spill and time since the spill. So we would expect to see the largest changes at the site number 3,4,5, somewhat less at 2,6,7, and much less at 1,8-11.

How do we measure changes among these areas to look at the effect of the spill?

First we estimate a model for the CPUE observed in the AFSC survey for each of 55 species for each year we have data (1984-2015; every 3 years, switch to every other year in 1999). We then generate a predictive density for each species on a 2x2km grid in each year. So we have a stack of 57 species distribution models for each location in each year. Using these predicted densities we can calculate all sorts of community metrics for each spatial location.

In this paper we focus on one metrics analyzed in several different ways:

· Total biomass

o By Taxonomy

§ Fish only (top 54 species)

§ Fish + common crabs (55 species)

§ Within Fish

· Sharks and allies vs. all others

o By life-habits

§ Pelagic vs. Benthic.

§ Functional feeding Groups

· there is some evidence that different groups have different exposure and susceptibility to oil.

After we have calculated a metric for each location, we can aggregate the predictions across all of the locations within a specified area to come up with a spatial mean or median for that metric in that year-area combination. I’d propose looking at each metric in three ways.

1. Treat each area as a univariate time-series for the metric.

a. Ask if there is a trend in the metric, or any notable changes at the dates that bracket the spill (1987 vs 1990 or 1993).

b. Ask if particular metrics return or diverge from pre-spill values

a. Any pattern with sites based on their proximity to the spill

2. Look at all areas simultaneously. Do the aggregate properties across all areas change with the spill

a. Is there a change in variance among sites

b. Does the pairwise correlation between site change in some way?

c. Are areas becoming more similar or more distinct post-spill?

**What is done so far:**

Work Completed (except where noted)

1. Identified the most common fish and mobile invert species

a. 52 fish

I. also did 6 species of abundant fish divided into “large” and “small” categories (20 cm break for all based on EcoPath model).

b. 3 crab

2. Fit several occurrence and abundance model for each.

a. Compared several models, picked a favorite.

I. Only used depth as a fixed covariate (bottom temperature proved inconvenient)

b. Saved model objects

3. Resampled from these estimated models to generate predictions for locations on a 2x2km grid in the Gulf of Alaska for each species (with uncertainty)

a. Made maps for each species

b. Made maps for aggregate quantities

i. Total biomass

1. All

2. Fish only

3. Cartilage vs. Boney

4. Pelagic vs. Benthic

5. To do:

a. Feeding categories.

b. Other functional traits.

4. Identified 11 areas with similar attributes to calculate index-standardized measures of abundance of each species.

a. Started creating univariate summaries of each site through time for aggregate measures.

b. Need to work up true portolio metrics

i. Variance, other metrics

c. To Do:

Assess whether we like the areas I chose or should add some.

**References**

Aydin, K., S. Gaichas, I. Ortiz, D. Kinzey, and N. Friday. 2007. A comparison of the Bering Sea, Gulf of Alaska, and Aleutian Islands large marine ecosystems through food web modeling. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-178, 298 p.

Stauffer G (2004) NOAA protocols for groundfish bottom trawl surveys of the nation’s fishery resources. NOAA Tech. Memo. NMFS-F/SPO-65. U.S. Department of Commerce, Washington, D.C., USA.

Armistead CE, Nichol DG (1993) Bottom trawl survey of the eastern Bering Sea continental shelf. NOAA Tech. Memo. NMFS AFSC-7. US Department of Commerce, Seattle, WA.

Doroff, A. M., A. R. DeGange, C. Lensink, B. E. Ballachey, J. L. Bodkin, and D. Bruden. 1993. Recovery of sea otter carcasses following the Exxon Valdez oil spill. Pages 285–288 in Exxon Valdez oil spill symposium abstracts. February 2–5, 1993, Anchorage, Alaska.

Thiago G. Martins, Daniel Simpson, Finn Lindgren and Havard Rue (2013)."Bayesian computing with INLA: New features.”

Piatt, J. F., and R. G. Ford. 1996. How many seabirds were killed by the Exxon Valdez oil spill? Pages 712-719 in S. D. Rice, R. B. Spies, D. A. Wolfe, and B. A. Wright editors. Proceedings of the Exxon Valdez Oil Spill Symposium, Anchorage, Alaska, 2-5 February, 1993. American Fish Society, Bethesda, MD.

Piatt, J.F., C.J. Lensink, W. Butler, M. Kendziorek, and D.R. Nysewander. 1990. Immediate impact of the 'Exxon Valdez' oil spill on marine birds. Auk 107: 387-397.