Using gradients in productivity to inform stock structure: a case study in Alaskan waters.

Kelli F. Johnson1\*, André E. Punt1, James T. Thorson2, and Kerim A. Aydin3

1School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, WA 98195-5020, USA

2Northwest Fisheries Science Center, National Marine Fisheries Service, NOAA, 2725 Montlake Blvd. East, Seattle, WA 98112-2097, USA

3Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA, 7600 Sand Point Way N.E., Seattle, WA 98115-6349, USA

\*Corresponding author: tel: +1 206 543 4270; fax: +1 206 685 7471; email: kfjohns@uw.edu

# Abstract

Todo: write abstract

# Keywords

Alaska, Atka mackerel, Gompertz model, Pacific cod, random field, stock structure, walleye pollock

# Introduction

The first step in the management of natural resources is to determine the management unit, which will depend on the management objectives. Thus, management units should evolve in concert with changing management objectives (Begg and Waldman, 1999). Instead, stock boundaries for many fish stocks remain unchanged over time despite pleas for better alignment of the spatial scale of management with the natural scale of targeted populations (Francis *et al*., 2007). Therefore, research and methods that define stock boundaries based on biology are becoming increasingly needed as managers shift their focus on single species maximum sustainble yield to optimal yield, which can incorporate biological, economic, and social objectives (Rothschild, 1973; Sissenwine, 1978).

Few marine fish species lack evidence for complex spatial structure (Conover *et al*., 2006; Cope and Punt, 2009). Furthermore, stock structures should not be seen as static; species ranges are predicted to shift with respect to depth and latitude as oceans warm in response to climate change (Perry *et al*., 2005; Dulvy *et al*., 2008; Nye *et al*., 2009), potentially increasing the mismatch between default, often political, and biological stock boundaries. Thus, accounting for spatial structure by properly identifying stock boundaries is necessary for several reasons, including but not limited to the i) fair allocation of catch among competing fisheries and/or nations (Kutkuhn, 1981; ), ii) protection of nursing and spawning areas (Smith *et al*., 1990; Stephenson, 1999), iii) development of optimal harvest and monitoring strategies (Begg *et al*., 1999a; Frank and Brickman, 2000), iv) reduction of localized depletion (todo:citation), v) proper identification of choke species (todo:citation), and vi) sustainability of prey resources (Bertrand *et al*., 2012).

Currently, no one method properly identifies all stocks. Thus, integrating the results from a suite of techniques to address multiple aspects of the biology and life-history characteristics of marine fishes increases the likelihood of detecting relevant variation among stocks (Hohn, 1997; Waldman *et al*., 1997; Coyle, 1998; Begg *et al*., 1999b; Abaunza *et al*., 2004; Hatfield *et al*., 2005). Several available techniques such as xxx require a substantive amount of costly data, necessitating a balance between accuracy and efficiency (Hilborn and Walters, 1992). Thus, techniques that use data routinely collected for assessment and management purposes (i.e. life-history parameters) are advantageous. As a bare minimum, catch data and distribution data from fishery-independent surveys, if available, can provide basic information on the areas used by a stock. Life-history data can provide inferences about genotypic variation resulting from environmental and anthropocentric influences by categorizing phenotypic expressions. In short, using life-history data to deliniate stock can provide a good starting point for more advanced stock identification techniques (e.g. biochemical, immunological, serological, behavioural, morphometric, meristic, mark-recapture, etc.), while being cost effective and efficient (Begg et al, 1999b). Unfortunately, fisheries data are subject to both observation and process error, which can complicate analyses.

State-space models can simultaneously estimate process and observation error making them ideal for ecological data. Their use in ecology began with time-series data for single populations (Staples, Taper, and Dennis, 2004; Dennis *et al*., 2006) and later extended into two-dimensional space for movement data (Jonsen, Myers, and Flemming, 2003) and multiple subpopulations (Ward *et al*., 2010) and species (Mattsson *et al*., 2013). Multivariate time-series analysis is one way to detect spatial patterns of correlation and synchrony across multiple locations (Ward *et al*., 2010) and can provide inferences on stock structure based on ecological differences. Unfortunately these models do not implicitly account for space.

Non-spatial and spatial models typically lead to different parameter estimates of density dependence and time series of population abundances (Thorson *et al*., In review), making it necessary to implicitly account for space. Accounting for space can be accomplished by replacing scalar functions with random field functions. Subsequently, spatially-explicit estimates of productivity and process error can be used to help inform stock structure. Todo: This paragraph needs more work and thought!

Here we quantify spatial differences in productivity for three species, walleye pollock (*Theragra chalcogramma*; hereafter referred to as pollock), Pacific cod (*Gadus macrocephalus*; hereafter referred to as cod), and Atka mackerel (*Pleurogrammus monopterygius*; hereafter referred to as mackerel) in the north east Pacific Ocean using spatially-explicit Gompertz population dynamics models, implemented within a state-space formulation. Pollock, cod, and mackerel were chosen because they are each important both ecologically and commercially in the Bering Sea Aleutian Island (BSAI) and Gulf of Alaska (GOA) large marine ecosystems. Pollock and cod represent the two largest fisheries by weight in all of Alaska (1,310,330 and 329,040 (t), respectively; st.nmfs.noaa.gov). Furthmore, pollock and mackerel are the main prey for cod, comprising approximately 19% and 27% of total cod stomach contents, respectively (Ortiz, 2007). Pollock and mackerel are also the two most common prey species (frequency of occurrence (FO) 46.4% and 39.6%, respectively) of the federally endangered Steller sea lion (*Eumetopias jubatus*; SSL), with cod being the fourth most common (FO = 16.1%) (Sinclair and Zeppelin, 2002). The method should be useful to researchers looking to establish stock structure hypotheses that will lead to robust ecosystem based fisheries management and minimize the chances of localized depletion.

# Methods

## Study area

The study area included the Aleutian management area (170° E to 170° W), a portion of the Southern Bering Sea management area (165° W to 170° W along the north side of the Aleutian Islands), and the Gulf of Alaska management area (132°40’ W to 170° W) (Figure 1). The AIs are an extensive archipelago comprised of 14 large and hundreds of small volcanic islands, characterised by a relatively narrow continental shelf, crossed by numerous deep passes. Conversely, the GOA is characterized by a semi-enclosed basin, with a relatively shallow shelf that varies in width from 5 km in the southeast to more than 200 km in the north.

## Data

Survey data were collected by the Alaska Fisheries Science Center of the National Marine Fisheries Service from 1982 to 2012 during the triennial GOA and AI bottom trawl surveys. These data are used to provide estimates of absolute or relative abundance (Aydin *et al*., 2013; A’mar *et al*., 2013), define distributions, and inform biological parameters of groundfish and crab to the the North Pacific Fishery Management Council (NPFMC).

Surveys employed a stratified random sampling design, with grids determined by NPFMC regulatory areas, which were further divided into depth strata. The AI surveys covered depths up to 500 m along the north side of the AIs from Unimak Pass (165° W) westward to Samalga Pass (170° W) and on both sides of the AIs from Samalga Pass to Stalemate Bank (170° E) (von Szalay *et al*., 2011). The GOA survey covered depths up to 1000 *m* from Samalga Pass to Dixon Entrance (54°30’ N, 132°40’ W) (Raring *et al*., 2011). Surveys in the AIs were conducted in 1983, 1986, 1991, 1994, 1997, 2000, 2002, 2004, 2006, 2010, and 2012. Surveys in the GOA were conducted in 1984, 1987, 1990, 1993, 1996, 1999, 2001, 2003, 2005, 2007, 2009, 2011, and 2013. The analysis was restricted to data collected after 1989 and east of 140° W because of the lack of standardized gear and effort measurements prior to 1990 (Munro and Hoff, 1995) and because of the Southeast Alaska Trawl Closure (Figure 1), respectively.

Tows of 15 min (estimated on-bottom duration) at speeds of 3 knots were conducted at each sampled station. For each tow, species composition of the catch was determined and species-specific catch per unit effort (*CPUE*) was calculated as the weight of the catch (kg) divided by the area swept (ha) for a given trawl (Figure 2). Global positioning (GPS) locations for each trawl were projected using the Albers equal-area conic projection, an equal area map projection.

## Spatial fields

Spatial data were defined as realizations of a stochastic process indexed by space

, Eq. 1

where the domain (Ɗ) is a fixed subset of ℝd (here d = 2). The point-referenced data of nonzero *CPUE* measurements, , where is a set indicating the spatial units at which the measurements were taken, were modelled using a stochastic partial differential equation (SPDE) approach proposed by Lindgren *et al*. (2011). Continuous spatial processes were represented using discretely indexed latent stationary Gaussian Markov random fields (GMRFs), i.e., the parameters of the process have a multivariate Normal distribution with mean **μ** = (μ­1, …, μn)′ and a spatially-structured covariance matrix Σ, whose generic element Σi,j = Cov(θi, θj) = . GMRFs use sparse precision matrices rather than dense covariance matrices, making them computationally efficient compared to Gaussian fields (Jona-Lasinio *et al*., 2012; Banerjee *et al*., 2004). Triangulation of Ɗ facilitates the transformation of a continuous space () to a discretely-indexed GMRF via a basis function

, Eq. 2

where *G* is the total number of vertices in the triangulation, { is the set of basis functions, and are the non-zero distributed weights. Basis functions were chosen such that was 1 at vertex *g* and 0 everywhere else (Lindgren *et al*., 2011). Here is the variance component for and is the (isotropic) Matérn spatial covariance function (Cressie, 1993), which depends on the Euclidean distance between the locations . It follows that,

, Eq. 3

where is the modified Bessel function of second kind and order, and measures the degree of smoothness. The range (), defined as the distance at which the spatial correlation becomes less than or equal to 0.1, was empirically derived using

, Eq. 4

where κ>0 is a scaling parameter and λ was set to 1.

## Gompertz model

Population dynamics for each species were modelled using a spatial equivalent of the log-linear stochastic Gompertz model,

, Eq. 5

where is the density at all spatial locations in year *t*, ρ is the strength of density dependence, Ω is a GMRF representing the spatial variation in productivity, and is a GMRF representing the spatial variation in process error in year *t*. It follows that, Ω ~ *MVN*(α**1**, ΣΩ) and Εt ~ *MVN*(0, ΣΕ), where α**1** is a vector and α is the mean productivity of the GMRF. ΣΩ and ΣΕ­ are defined as above (Eq. 3) with independent variance components and equal Matérn spatial covariance functions.

Initial conditions were specified using the following equation:

, Eq. 6

where ϕ**1** is the log-ratio of expected abundance in the initial year (*t*=1) and the median of the stationary distribution for abundance at equilibrium.

## Estimation procedure

Data were assumed to follow a lognormal distribution

, Eq. 7

where, is the predicted density at a given vertex *i* located at the points [*x*, y] on the triangulation network, is the standard deviation of the observation error, and are the *CPUE* observations located in the triangulation vertex *i* in year *t*. and are statistically-independent of and conditional on a fixed value for and , implying partial separability.

GMRF matrices were computed using functions available in the INLA package (Rue *et al*., 2014; Illian *et al*., 2012) implemented in the R statistical software environment (R Core Team, 2014). A single triangulation network used to facilitate the comparison of productivity across species. Parameters of the models were estimated using the Template Model Builder (TMB) software (Kristensen *et al*., 2013). TMB uses the integrated nested Laplace approximation (Skaug and Fournier, 2006) to calculate the marginal likelihood of the fixed-effect parameters integrated across all random fields, while also calculating the gradient of the marginal likelihood via automatic differentiation. Fixed effects are then subsequently optimized using traditional tools in R by maximizing the log-marginal likelihood. Confidence intervals are available using the hessian matrix and delta-method.

# Results

Todo: Write results

# Discussion

Todo: Write discussion

# Acknowledgements

Todo: Write acknowledgments

# References

Todo: Add remaining references

Aydin, K., Barbeaux, S. J., Conners, M. E., Conrath, C., Dalton, M., DiCosimo, J., Echave, K., Hanselman, D., Hoff, J., Honkalehto, T., Hulson, P. J., Ianelli, J., Kotwicki, S., Lowe, S., Lunsford, C., McKelvey, D., Nichol, D., Ormseth, O. A., Palsson, W., Rodgveller, C. J., Rooper, C. N., Spencer, P., Spies, I., Stockhausen, W., TenBrink, T., Thompson, G., Tribuzio, C., Wilderbuer, T., Williamson, N. 2013. Stock assessment and fishery evaluation report for the groundfish resources of the Bering Sea/Aleutian Islands region. North Pacific Fisheries Management Council, Anchorage, AK.

A’mar, T., Aydin, K., Conners, M. E., Conrath, C., Dalton, M., Davis, O., Dorn, M., Echave, K., Friday, N., Green, K., Hanselman, D., Heifetz, J., Hulson, P. J., Ianelli, J., Jones, D., Jaenicke, M., Lowe, S., Lunsford, C., Meyer, S., McGilliard, C., Nichol, D., Ormseth, O. A., Palsson, W., Rodgveller, C. J., Rumble, J., Shotwell, K., Slater, L., Spalinger, K., Spencer, P., Spies, I., Stewart, I., Stichart, M., Stockhausen, W., Stram, D., TenBrink, T., Tribuzio, C., Turnock, J. 2013. Stock assessment and fishery evaluation report for the groundfish resources of the Gulf of Alaska. North Pacific Fisheries Management Council, Anchorage, AK.

Banerjee, S., Carlin, B. P., Gelfand, A. E. 2004. Hierarchical Modeling and Analysis for Spatial Data. Crc Press.

Cressie, N. A. C. 1993. Statistics for Spatial Data. Wiley, NY.

Illian, J. B., Sørbye, S. H., Rue, H. 2012. A toolbox for fitting complex spatial point process models using integrated nested Laplace approximation (INLA). The Annals of Applied Statistics, 6:1499-1530.

Jona-Lasinio, G., Mastrantonio, G. Pollice, A. 2012. Discussing the “big n problem”. Statistical Methods and Applications, 22(1):97-112.

Kristensen, K. Thygesen, U. H., Andersen, K. H., Beyer, J. E. 2013. Estimating spatio-temporal dynamics of size-structured populations. Canadian Journal of Fisheries and Aquatic Sciences, 71:326-336.

Lindgren, F., Rue, H., Lindstrӧm, J. 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach [with discussion]. Journal of the Royal Statistical Society B, 73(4):423-498.

Munro, P. T., Hoff, R. Z. 1995. Two demersal trawl surveys in the Gulf of Alaska: implications of survey design and methods. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-50, 139 p.

R Core Team. 2014. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Raring, N. W., von Szalay, P. G., Shaw, F. R., Wilkins, M. E. 2011. Data Report: 2001 Gulf of Alaska bottom trawl survey. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-225, 179 p.

Rue, H., Martino, S., Lindgren, F., Simpson, D., Riebler, A. 2014. INLA: functions which allow to perform full Bayesian analysis of latent Gaussian models using Integrated Nested Laplace Approximation. R package version 0.0-1392038736.

Skaug, H., Fournier, D. 2006. Automatic approximation of the marginal likelihood in non-Gaussian hierarchical models. Computational Statistics & Data Analysis, 51:699-709.

von Szalay, P. G., Rooper, C. N., Raring, N. W., Martin, M. H. 2011. Data Report: 2010 Aleutian Islands bottom trawl survey. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-215, 153 p.

Ward, E. J., Chirakkal, H., González-Suárez, M., Aurioles-Gamboa, D., Holmes, E. E., Gerber, L. 2010. Inferring spatial structure from time-series data: using multivariate state-space models to detect metapopulation structure of California sea lions in the Gulf of California, Mexico. Journal of Applied Ecology, 47: 47–56.

Wilkins, M. 2009. Bottom trawl survey of groundfish resources in the Aleutian Islands region. Alaska Fisheries Science Center, Seattle, WA.

# Tables

Todo: result tables

# Figures

Todo: Add remaining figures

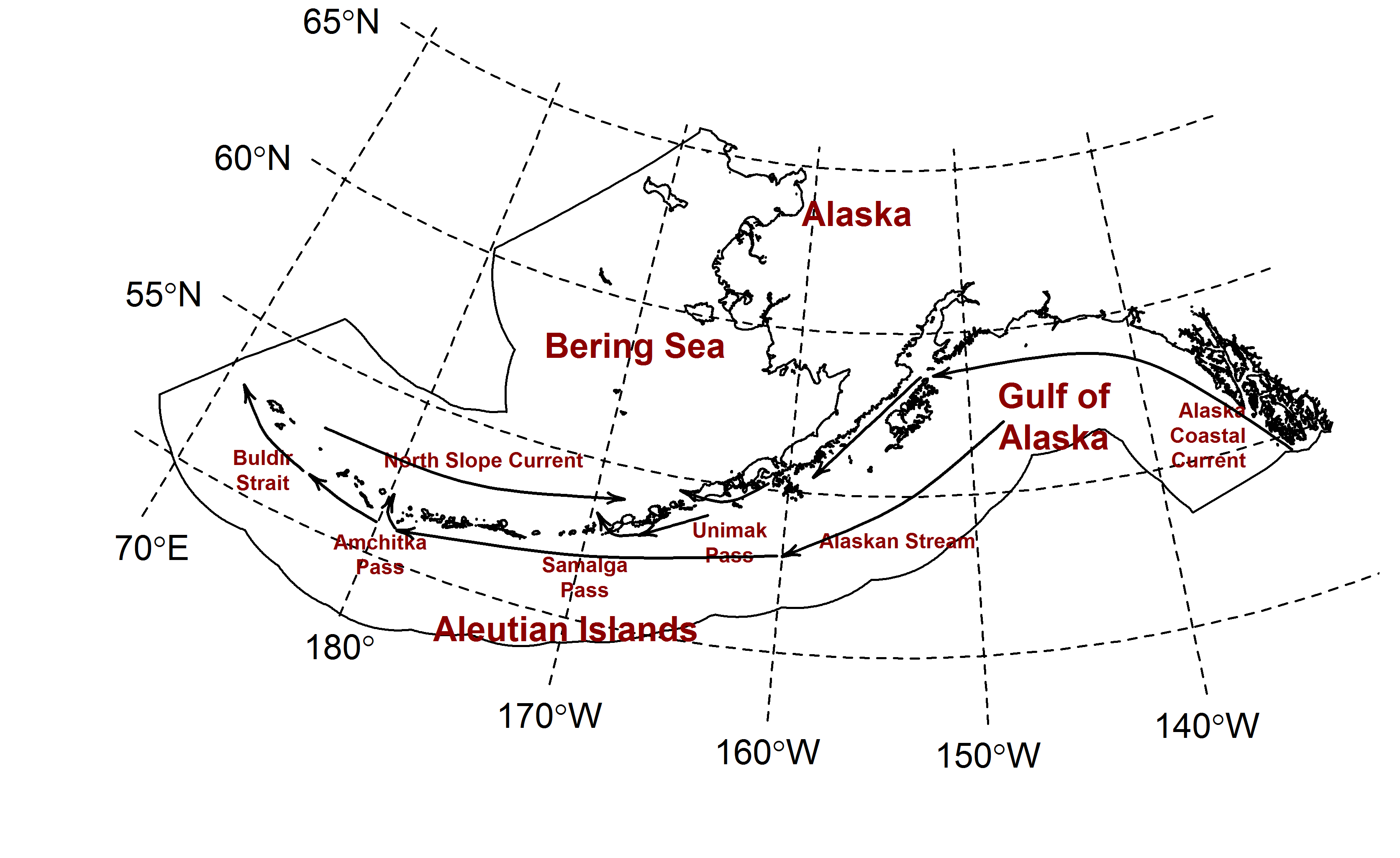


Figure 1. Map of the study area with major currents. Black outline marks the Alaskan coast and the Exclusive Economic Zone. Todo: Add Southeast Alaska Trawl Closure to the map and fix the axis label for 170E.

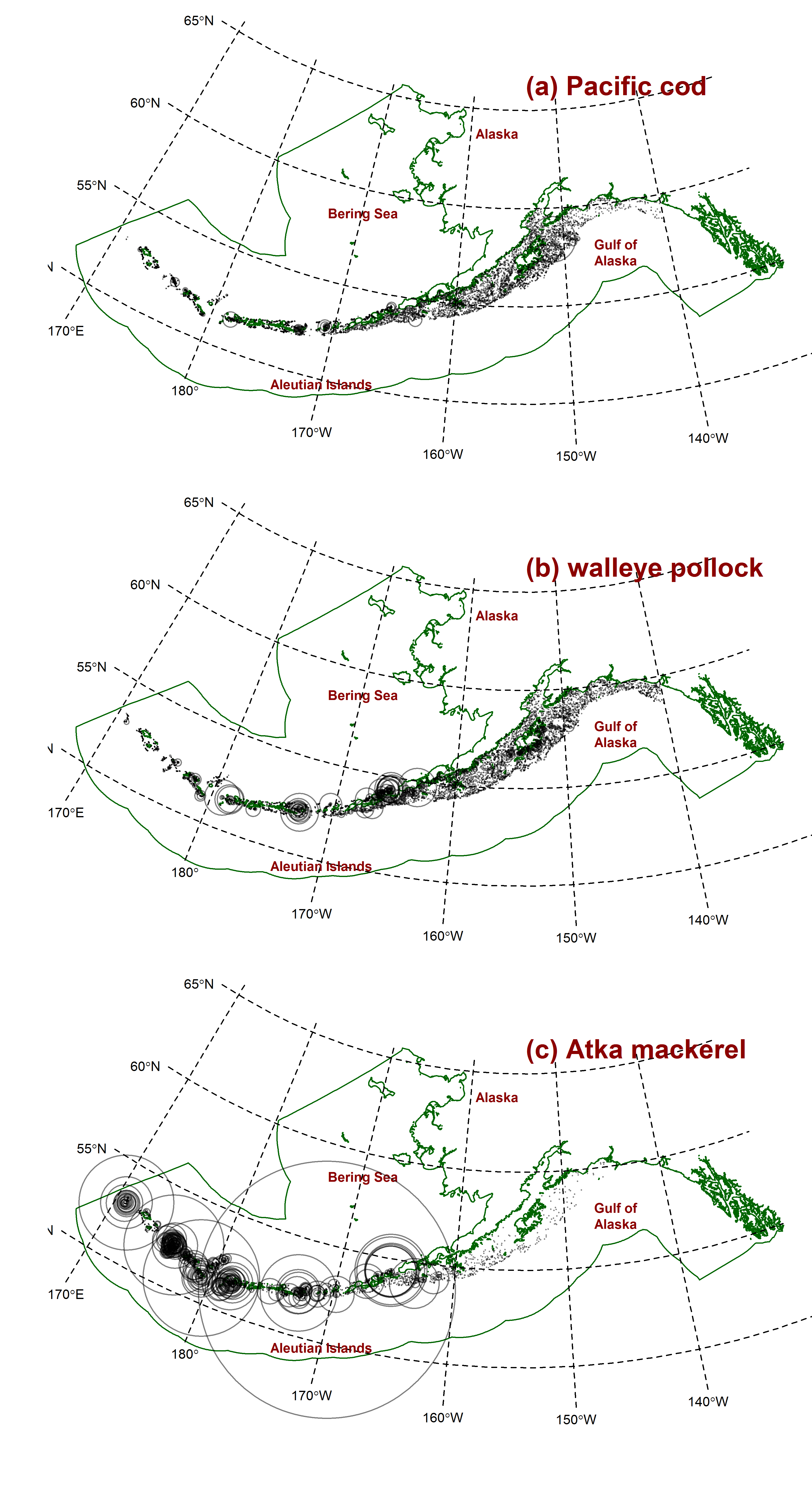


Figure 2. Log catch per unit effort (CPUE, kg ha-1) for (a)Pacific cod, (b) walleye pollock, and (c) Atka mackerel for all years (1990-2013) where the size of the circle at each sampled location is proportional to the observed CPUE. Green line marks the Alaskan coast and the Exclusive Economic Zone.