**Does incorporating spatio-temporal correlations among fishes and biogenic habitat improve estimates of abundance trends and distribution shifts?**

James T. Thorson1, Lewis A.K. Barnett2

1 Fisheries Resource Assessment and Monitoring Division, Northwest Fisheries Science Center, National Marine Fisheries Service, NOAA, Seattle, WA, USA

2 School of Aquatic & Fishery Sciences, University of Washington, Seattle, WA, USA; under contract to Fishery Resource Analysis and Monitoring Division, Northwest Fisheries Science Center, NOAA, Seattle, WA, USA

**keywords**: Joint dynamic species distribution model (JDSDM); structure-forming invertebrates; spatio-temporal model; Pacific rockfish; index standardization; fish distribution shift

**Abstract**

Several approaches have been developed over the last decade to simultaneously estimate distribution or density for multiple species (e.g., “joint species distribution” or “multi-species occupancy” models). However, there has been little research comparing estimates of abundance trends or distribution shifts from these multispecies models with similar single-species estimates. We therefore seek to determine whether a model including correlations among species (and particularly species that may affect habitat quality, termed “biogenic habitat”) improves predictive performance or decreases standard errors for estimates of total biomass and distribution shift relative to similar single-species models. To accomplish this objective, we apply a vector-autoregressive spatio-temporal (VAST) model that simultaneously estimates spatio-temporal variation in density for multiple species, and present an application of this model using data for eight US Pacific Coast rockfishes (*Sebastes* spp.), thornyheads (*Sebastolobus* spp.), and structure-forming invertebrates (SFI). We identified three fish groups having similar spatial distribution (northern *Sebastes*, coastwide *Sebastes*, and *Sebastolobus* species), and estimate differences among groups in their association with SFI. The multispecies model was more parsimonious and had better predictive performance than fitting a single-species model to each taxon individually, and estimated fine-scale variation in density even for species with relatively few encounters (which the single-species model was unable to do). However, the single-species models estimated similar abundance trends and distribution shifts to those of the multi-species model but with smaller standard errors. We therefore conclude that spatial variation in density (and annual variation in these patterns) are correlated among fishes and SFI, with congeneric fishes more correlated than species from different genera. However, explicitly modelling correlations among fishes and biogenic habitat does not seem to improve estimates of abundance trends or distribution shifts for these fishes.

**Introduction**

There are several benefits to simultaneously analysing the distribution and density of multiple species within a natural community. Multispecies models of spatial distribution can estimate associations among species (Thorson et al. In press-b, Latimer et al. 2009, Ovaskainen et al. 2016), such that the presence or absence of a given species can be used as an indicator of habitat for other species when reliable habitat variables are otherwise lacking (Ovaskainen et al. 2010). Multispecies models fitted to presence/absence data (termed “multispecies occupancy models”) can also be used in some cases to identify the impact of management actions more efficiently than using single-species occupancy models (Zipkin et al. 2010). Furthermore, research shows that estimating the distribution for each species individually and then summarizing community-level properties by stacking results from single-species analyses can result in improper inference about ecological communities (Clark et al. 2014).

The predictive performance of species distribution models is often improved when available covariates are included that are informative about habitat quality. Unfortunately, environmental variables associated with habitat quality are difficult to measure for many species, including demersal marine fishes. To overcome this difficulty, new species distribution modelling (SDM) techniques may allow differences in habitat to be inferred from spatial variation in the density of species with similar habitat requirements (Latimer et al. 2009, Ovaskainen et al. 2010). For example, joint species distribution models have previously been used to show strong covariation in population density among US Pacific Coast rockfishes and thornyheads (*Sebastes* and *Sebastolobus* spp.), and these correlations imply that population density of one species is informative about the density of correlated species (Thorson et al. 2015a). Similarly, joint dynamic species distribution models (JDSDMs) can estimate abundance trends for infrequently encountered species, and have revealed similarities in spatio-temporal dynamics among related butterfly species (Thorson et al. In press-b). However, JDSDMs have not previously been used to explore associations between fishes and species that are associated with specific habitat features (e.g., structure-forming invertebrates, SFI).

Marine fishes are intensively managed in many parts of the developed world, and their management of marine fisheries is strongly linked to estimates of population status and productivity from population models (termed “stock assessment models”) throughout North America and Europe (Methot 2009, Maunder and Punt 2013). Although these stock assessment models often integrate many different types of information, time series that are proportional to population abundance (“abundance indices”) are often among the most critical (Francis 2011). For this reason, there is considerable research regarding best-practices for minimizing error when estimating abundance indices for fishes from survey data (Walters 2003, Maunder and Punt 2004, Shelton et al. 2014). Similarly, survey data are increasingly used to estimate shifts in fish distribution over time (e.g., due to climate change), and distribution shifts are often measured by estimating the centroid of the population’s distribution and shifts in this centroid over time (Perry et al. 2005, Pinsky et al. 2013). Research suggests that spatio-temporal models are statistically efficient and can improve precision for estimates of abundance indices or distribution shifts relative to nonspatial models given limited available data (Thorson et al. 2015b, 2016). Recently, novel methods have been proposed for estimating abundance indices by simultaneously fitting a JDSDM to data for multiple species (Thorson et al. In press-a). However, there is little research comparing the single- and multi-species approaches to estimating abundance indices for marine fishes.

For three reasons, Pacific rockfishes and their close relatives provide an interesting example when studying associations between fishes and species that affect habitat suitability (“biogenic habitat”), or the potential benefit of these associations when estimating abundance indices or distribution shifts. Most importantly, Pacific rockfishes manifest an astounding diversity of species, with more than 65 species co-occurring in the northeast Pacific (Hyde and Vetter 2007), and exhibit a wide range of life-history strategies (Love et al. 2002, Mangel et al. 2007). Given this life-history diversity, rockfishes likely include species whose spatial distributions are both strongly correlated and relatively uncorrelated with SFI. Second, Pacific rockfishes differ in functional traits related to feeding type and efficiency (eye and gill raker size, Ingram and Shurin (2009)), so species with similar spatial distribution and feeding types might exhibit correlated changes in productivity over time in response to variable food supply. Therefore, bottom-up drivers abundance or distribution changes would result in correlated abundance or distribution changes over time for species with similar feeding types. Third, many Pacific rockfishes have low and extremely variable population densities (Thorson et al. 2011), such that single-species estimates of trends in population abundance or population distribution are frequently imprecise (Thorson et al. 2015b, 2016). Given these characteristics of the rockfish assemblage, including information about species associations and biogenic habitat when estimating population abundance may increase precision and thereby improve stock assessments.

Given the potential benefit of estimating habitat quality from the density of co-occurring marine species when estimating abundance indices, we seek to simultaneously estimate the density of Pacific rockfishes and structure-forming invertebrates at a coastwide scale. Specifically, we seek to answer three questions: (1) do Pacific rockfishes have an association with structure-forming invertebrates on the US West Coast? (2) Is this association similar or variable among rockfish species? and (3) Does the inclusion of information regarding co-occurrence (either among rockfishes, or between rockfish and SFI) improve estimates of rockfish abundance trends or population distribution? To address these questions, we develop a vector-autoregressive spatio-temporal (VAST) model for jointly analysing catch-rate data for fish and structure-forming invertebrates, and apply the model to data for eight rockfish species and SFI from 2003-2014.

**Methods**

**Pacific rockfishes**

Pacific rockfishes, genus *Sebastes*, and thornyheads, genus *Sebastolobus*, (hereafter collectively called “rockfishes”) are one of the dominant species-groups within the assemblage of bottom-associated fishes off the US West Coast. Pacific rockfishes in this region are monitored by the West Coast groundfish bottom trawl survey (WCGBTS) conducted annually by the Northwest Fisheries Science Center since 2003 (Bradburn et al. 2011). The WCGBTS covers areas from the Canada to Mexico border between 55 to 1280 m depth, and survey stations for each year are chosen at random within strata defined by depth and latitude (two regions divided at Point Conception, CA). Four commercial vessels (20-28m length) are chartered each year to sample from mid-May to late October, conducting approximately 15 min tows at a speed of 2 knots using a standard Aberdeen-type trawl with 3.8cm mesh codend liner, 25.9m headrope, and 31.7m footrope. All fishes and invertebrates are sorted at sea to the lowest possible taxon, and their wet weight is measured. For the purposes of our analysis, we take the midpoint of each haul to represent the location of each biological sample.

We analyse these survey data between the years 2003 and 2014, focusing on structure-forming invertebrates and eight species of Pacific rockfish (Table 1) that are frequently captured within the survey and for which there was previous documentation of association with structure-forming invertebrates at fine spatial scales (Love et al. 2002). We aggregate the structure-forming invertebrate taxa into a single grouping to obtain adequate encounter rates for estimating the distribution for structure-forming invertebrates. This SFI group primarily consists of sponges (phylum Porifera), anemones (order Actinaria), and sea pens (order Pennatulacea), along with fewer observations of true corals (subclass Hexacorallia), and other soft corals (subclass Octocorallia). Although the survey is primarily designed to capture demersal fishes and is not as effective as visual methods for assessing structure-forming invertebrates, it is the primary source of data for estimating spatio-temporal associations between demersal fishes and biogenic habitat at large spatial and temporal scales off the US West Coast. Bottom-trawl samples have been shown to be a good predictor of biogenic habitat distribution in areas such as the eastern Bering Sea based on validation using camera surveys (Rooper et al. 2016).

**Vector-autoregressive spatio-temporal (VAST) model**

We seek to estimate the association among fishes and structure-forming invertebrates, and therefore model correlations among density for each taxon (indicating fish species or the SFI group) at location and time . To do so, we build upon recent research regarding joint dynamic species distribution models (JDSDM). In particular, we propose a vector autoregressive spatio-temporal (VAST) model, where the probability distribution for catch data for each taxon is decomposed into two components, representing (1) the probability of encounter for taxon in the th sample, and (2) the expected catch rate given that taxon is encountered (Maunder and Punt 2004, Martin et al. 2005):

where is the Dirac delta-function, is a lognormal probability distribution function for value given a log-mean of and a variance of , is the location of the th sample, and is the year for the th sample. Encounter probability is approximated using a logit-linked linear predictor

where is an intercept for encounter probability each taxon and time , is spatio-temporal variation in log-expected density, and is a “vessel effect” for the vessel conducting the th sample. Vessel effects are included because the WCGBTS is obtained using 3-4 different vessels per year, and previous research indicates that vessels in each year have small but important variation in fishing behaviour and resulting catch rates (Helser et al. 2004, Thorson and Ward 2014). Expected catch rates when a species is encountered are similarly approximated using a log-linked linear predictor:

where is an intercept, is spatio-temporal variation, and is a vessel effect.

The VAST model that we developed involves specifying a probability distribution for spatio-temporal variation ( and ) and vessel effects ( and ). For each modelled year, we therefore specify a three-dimensional Gaussian process for spatio-temporal variation:

where is a matrix composed of at every modelled location and taxon in year , is the correlation among locations, and is the covariance in spatio-temporal variation among species (where is follows an identical distribution but with and in place of and ). Spatial correlation between location and location follows a Matérn correlation function given the distance between the two:

where is smoothness (fixed at 1.0), is the Bessel function, and governs the distance over which locations are uncorrelated (and where is defined identically but with in place of ). Covariance among species is modelled using a factor-analysis decomposition:

where is a by matrix defining the first columns of the Cholesky decomposition of covariance matrix , and is defined identically but with in place of (see Thorson et al. (In press-b) and Warton et al. (2015) for details regarding this factor-analysis decomposition). We also specify a factor-analysis decomposition for the covariance among vessel effects:

where is the vector of vessel effects for the th vessel for each species (and where is defined identically but with in place of ). This factor-analysis decomposition allows the analyst to select an appropriate number of factors for approximating spatio-temporal covariation or covariation among vessels, where . In the following, however, we specify full rank for each covariance (i.e., ).

Parameters are estimated for this VAST model by maximizing the marginal likelihood of fixed effects given available data. We treat the intercept parameters for each species ( and ), the spatial scale of spatio-temporal variation ( and ), the covariation among species ( and ), the covariation among vessels ( and ), and the magnitude of residual variation in positive catch rates for each species () as fixed effects. We calculate the marginal likelihood of fixed effects while integrating across random effects (and for each year , and and for every vessel ), and use the Laplace approximation to approximate this multidimensional integral (Skaug and Fournier 2006). The Laplace approximation is implemented using Template Model Builder (Kristensen et al. 2016), and Template Model Builder also provides the gradient of the approximated marginal likelihood with respect to all fixed effects. The maximum-likelihood estimate (MLE) of fixed effects is identified using a gradient-based nonlinear minimizer within the R statistical environment (R Core Team 2015). We use Revolution Open R for low-level parallelization of all computations (<http://www.revolutionanalytics.com/revolution-r-open>), and we distribute code for applying the VAST model to other data sets as an R package on the author’s website ([www.github.com/james-thorson/VAST](http://www.github.com/james-thorson/VAST)). We have confirmed that the VAST model provides identical parameter estimates to a previous spatio-temporal index standardization model, (SpatialDeltaGLMM, Thorson et al. 2015b), when applied to data for a single species. However, the VAST model also incorporates most capabilities of spatial dynamic factor analysis for monitoring trends in community abundance or conducting species ordination (Thorson et al. In press-b).

After parameters are estimated, we estimate total biomass for each species in each year (an “index of abundance”) as well as the centroid of the distribution for each species in each year (termed “center of gravity”) using standard formulae (Thorson et al. 2015b, 2016). We compare model performance when fitting all species simultaneously (the “multi-species analysis”) to a conventional “single-species analysis” where each species is fitted individually using the VAST model. To compare performance between single- and multi-species models, we compute the Akaike information criterion, AIC (Akaike 1974), as a measure of parsimony for each model while computing the “single-species” AIC as the sum of the AIC for the VAST model fitted to each individual species. We also conduct a 10-fold crossvalidation analysis to determine whether multi- or single-species analyses have greater predictive ability. To do so, we divide the data into 10 similarly-sized partitions. For the first cross-validation, we estimate model parameters only using data in partitions 2-10, and calculate the probability of data in partition 1 using the predictive distribution given estimated parameters. This process is repeated for all 10 partitions for the multi-species model. For the single-species model, we conduct this 10-fold cross-validation for each species individually, and then sum the resulting log-probabilities for each species.

**Results**

Inspection of density maps for eight fishes and structure-forming invertebrates shows that species are unevenly distributed throughout the California Current (Fig. 1). By distribution, the fishes can be broadly classified into three groups: coastwide *Sebastes* spp. (splitnose, stripetail, and greenspotted); northern *Sebastes* spp. (POP, sharpchin, and darkblotched); and *Sebastolobus* spp. (longspine and shortspine thornyheads). The thornyheads are distinguished by having increased densities in the deepest waters furthest from the US coast. Structure-forming invertebrates are found at highest densities offshore near northern Oregon, close to the Oregon-California border and offshore from south of Monterey Bay through the Southern California Bight.

Our species classifications are supported by the estimated covariance matrices (Fig. 2), where longspine and shortspine thornyheads have high pairwise correlations in both encounter probability and positive catch rates (0.5-0.7). Encounter probability of longspine is negatively associated with encounter probability of *Sebastes* spp., while encounter probability of shortspine has both positive and negative associations with different *Sebastes* spp. This difference between thornyhead species is also apparent in distribution maps (Fig. 1), where longspine has the deepest distribution of any fish in our analysis, whereas shortspine occupies a more shoreward distribution that overlaps with the spatial distribution for several *Sebastes* (e.g., darkblotched, POP, and splitnose). The *Sebastes* spp. all generally have high correlations (0.5-0.9) with one another for encounter probability (; Fig. 2 top panel), but the northern vs. coastwide groups are strongly distinguished by correlations in positive catch rates (; Fig. 2 bottom panel), where darkblotched, POP, and sharpchin have higher correlations with one-another (0.9) than with splitnose, stripetail, and greenspotted (0.2-0.7). At this coastwide spatial scale, *Sebastolobus* spp. generally have increased encounter probability when structure-forming invertebrates are found, whereas coastwide and northern rockfish groups have somewhat decreased encounter probability in these cases. When fishes and SFI are encountered, however, an increased catch of SFI is associated with increased catch for all fishes except stripetail and sharpchin rockfish.

We next compare estimates of biomass trends using multi-species and single-species estimates (Fig. 3). Biomass trends are broadly similar between models, and particularly for SFI, which shows a trend of increased biomass since 2008. Biomass trend estimates are most different between multi- and single-species models for the group of northern *Sebastes* spp. (darkblotched, sharpchin, and POP; Fig. 3 middle row). For example, the multispecies model estimates lower abundance for POP in 2008 than the single-species model. This lower estimate for POP in 2008 using the multispecies model reflects a similar decrease in abundance for darkblotched rockfish in 2008 using the multispecies model – the estimate for POP in this year is “shrunk” towards the estimate for darkblotched rockfish.

Estimates of variation and trends in center-of-gravity (COG) are also generally similar between multi- and single-species model outputs (Fig. 4). The notable exceptions are again the northern *Sebastes* spp., specifically POP and sharpchin rockfish, which both have relatively few encounters relative to other species (around 500 each, see Table 1). For POP and sharpchin, the single-species estimates of COG are nearly 100 km further south than COG estimates from the multispecies model (Fig. 4 middle row). By sharing information about positive densities (as shown in lower panel of Fig. 2), the multispecies model estimates greater variation in density for these species between different locations off Oregon and Washington, and therefore estimates a more northward distribution than the single-species model for POP and sharpchin (Appendix A and Fig. S1). We again interpret this as a consequence of statistical “shrinkage” for these species, where the multispecies model is sharing information among northern *Sebastes* spp. to infer density hotspots.

Finally, a comparison of standard errors (Fig. 5) shows that the multispecies model generally has wider confidence intervals for estimating biomass indices and center-of-gravity. This increased standard error presumably occurs because the multispecies model estimates greater spatial variation in density (Fig. S1). For POP for example, the single-species model estimates little spatial pattern except an increase in density moving northward along the coast, while the multi-species model estimates density hotspots in the same mid-depth areas off the Washington coast as it estimates as good habitat for splitnose and darkblotched rockfishes (Fig. S1).

Despite estimating wider standard errors for abundance indices and distribution shifts, the multispecies model provides a more accurate and parsimonious fit to available data. The multispecies model has an Akaike Information Criterion (AIC) score that is 5692.0 better than the combined AIC for single-species models, despite the multispecies model estimating an additional 112 parameters (409 fixed effects for the multi-species vs. 297 total among all single-species models). The improvement in fit for the multispecies model is also supported by the 10-fold crossvalidation analysis, where the multi-species model has a 4-5% greater predictive probability than when analysing each species individually (Table 2). This improvement in predictive score presumably arises because the multispecies model identifies fine-scale differences in species density for all taxa, and these fine-scale density estimates are on average a useful prediction of density variation.

**Discussion**

We have used a joint dynamics species distribution model (JDSDM) to illustrate strong associations (both positive and negative) between deep-water demersal fishes and structure-forming invertebrates at broad spatial scales along the US portion of the California Current. These associations vary substantially between two genera, *Sebastolobus* (thornyheads) and *Sebastes*, where *Sebastes* can be further divided into northern and coastwide species. Previous work has shown phylogenetic signals in covariation among fishes (Thorson et al. In press-b, 2015a) or other species (Ovaskainen et al. 2010), but ours is the first study (1) to use a spatio-temporal statistical model to estimate covariance between fishes and structure-forming invertebrates, and (2) to decompose this covariation into components representing encounter probabilities vs. positive catch rates (i.e., using the delta-models that are conventional in fisheries science; Maunder and Punt (2004)). Although the JDSDM provided a more parsimonious fit to available data than single-species models (as shown by AIC and crossvalidation analysis), the multi-species analysis resulted in slightly-wider confidence interval estimates than analysing data for each species individually.

At a coastwide spatial scale, we estimate an increased encounter probability for *Sebastolobus* and a decreased encounter probability for *Sebastes* species where SFI are present. By contrast, alternative visual sampling at fine spatial scales often shows a large increase in *Sebastes* density given the presence of SFI, and *Sebastolobus* densities are less often reported to be associated with biogenic habitat (Brodeur 2001, Tissot et al. 2008, Yoklavich and O’Connell 2008, du Preez and Tunnicliffe 2011). Recent research suggests that correlations in distribution among species will often differ when looking at small and large scales (Ovaskainen et al. 2016), and this may explain why our results differ from those from fine-scale visual sampling. Alternatively, differences in results may arise because visual sampling often occurs in rocky habitats, whereas our analysis relies upon bottom trawl data that are primarily available in soft-sediment habitats. We recommend future research combining data from small and coastwide scales (and both hard- and soft-bottom habitat) within a single spatio-temporal statistical model, where fine scales could be informed by either fishery-dependent catch rate data or direct observations (Thorson et al. In press-a, Jagielo et al. 2003, Shelton et al. 2014, Rooper et al. 2016). We also recommend future research to include habitat variables and associations within size-structured spatio-temporal models (e.g., Kristensen et al. 2014, Nielsen et al. 2014). These models could then estimate separate habitat associations for juvenile and adult fishes, and be used to target spatial management towards the more vulnerable or sensitive life stage for protected species.

Based on our results, we find that simultaneously modelling fishes and SFI yields more precise and parsimonious predictions of density, and also facilitates estimating variation in density at finer spatial scales than single-species models, even for species with few encounters (e.g., POP and sharpchin rockfishes). However, incorporating these associations when estimating trends in abundance or distribution does not shrink confidence intervals. For an ecologist conducting a stock assessment, incorporating multispecies data may complicate their description of estimated abundance indices, thereby decreasing stakeholder trust in the stock assessment process. We therefore imagine that our results will encourage many assessment scientists to continue using single-species models. From a broader perspective, however, the increased parsimony and out-of-sample predictive ability of the multispecies model indicates that estimates of local density are generally improved by jointly modelling multiple species (including both fishes and biogenic habitat). These estimates of local density could be used to prioritize areas for spatial management that have high density of structure-forming invertebrates and fishes. We therefore suggest further research regarding the association of fished species and biogenic habitat, including the likely impact of spatial management on fishery productivity in the West Coast groundfish fishery.

**Acknowledgements**

We thank NWFSC FRAM Fisheries Research Survey Team and fishermen crew on the US West Coast Groundfish Bottom Trawl Survey for collecting the data, and we thank Michelle McClure, Trevor Branch, and Tim Essington for helpful comments and discussion. LAKB gratefully acknowledges funding from the Joint Institute for the Study of the Atmosphere and Ocean (JISAO) under NOAA Cooperative Agreement No. NA15OAR4320063, Contribution No. 2016-01-38. We also thank Mary Yoklavich, Joe Bizzarro, and Chris Rooper for comments on an earlier draft.

**Bibliography**

Akaike, H. 1974. New look at statistical-model identification. IEEE Trans. Autom. Control **AC19**(6): 716–723.

Bradburn, M.J., Keller, A.A., and Horness, B.H. 2011. The 2003 to 2008 US West Coast bottom trawl surveys of groundfish resources off Washington, Oregon, and California: estimates of distribution, abundance, length, and age composition. NOAA Tech. Memo., US Department of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Northwest Fisheries Science Center, Seattle, WA.

Brodeur, R.D. 2001. Habitat-specific distribution of Pacific ocean perch (Sebastes alutus) in Pribilof Canyon, Bering Sea. Cont. Shelf Res. **21**(3): 207–224. doi:10.1016/S0278-4343(00)00083-2.

Clark, J.S., Gelfand, A.E., Woodall, C.W., and Zhu, K. 2014. More than the sum of the parts: forest climate response from joint species distribution models. Ecol. Appl. **24**(5): 990–999. doi:10.1890/13-1015.1.

Francis, R.I.C.C. 2011. Data weighting in statistical fisheries stock assessment models. Can. J. Fish. Aquat. Sci. **68**(6): 1124–1138.

Helser, T.E., Punt, A.E., and Methot, R.D. 2004. A generalized linear mixed model analysis of a multi-vessel fishery resource survey. Fish. Res. **70**(2-3): 251–264. doi:10.1016/j.fishres.2004.08.007.

Hyde, J.R., and Vetter, R.D. 2007. The origin, evolution, and diversification of rockfishes of the genus Sebastes (Cuvier). Mol. Phylogenet. Evol. **44**(2): 790–811.

Ingram, T., and Shurin, J.B. 2009. Trait-based assembly and phylogenetic structure in northeast Pacific rockfish assemblages. Ecology **90**(9): 2444–2453.

Jagielo, T., Hoffmann, A., Tagart, J., and Zimmermann, M. 2003. Demersal groundfish densities in trawlable and untrawlable habitats off Washington: implications for the estimation of habitat bias in trawl surveys. Fish. Bull. **101**(3): 545–565.

Kristensen, K., Nielsen, A., Berg, C.W., Skaug, H., and Bell, B.M. 2016. TMB: Automatic Differentiation and Laplace Approximation. J. Stat. Softw. **70**(5): 1–21. doi:doi: 10.18637/jss.v070.i05.

Kristensen, K., Thygesen, U.H., Andersen, K.H., and Beyer, J.E. 2014. Estimating spatio-temporal dynamics of size-structured populations. Can. J. Fish. Aquat. Sci. **71**(2): 326–336. doi:10.1139/cjfas-2013-0151.

Latimer, A.M., Banerjee, S., Sang Jr, H., Mosher, E.S., and Silander Jr, J.A. 2009. Hierarchical models facilitate spatial analysis of large data sets: a case study on invasive plant species in the northeastern United States. Ecol. Lett. **12**(2): 144–154.

Love, M.S., Yoklavich, M.M., and Thorsteinson, L.K. 2002. The rockfishes of the northeast Pacific. University of California Press.

Mangel, M., Kindsvater, H.K., and Bonsall, M.B. 2007. Evolutionary analysis of life span, competition, and adaptive radiation, motivated by the Pacific rockfishes (Sebastes). Evolution **61**(5): 1208–1224.

Martin, T.G., Wintle, B.A., Rhodes, J.R., Kuhnert, P.M., Field, S.A., Low-Choy, S.J., Tyre, A.J., and Possingham, H.P. 2005. Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. Ecol. Lett. **8**(11): 1235–1246.

Maunder, M.N., and Punt, A.E. 2004. Standardizing catch and effort data: a review of recent approaches. Fish. Res. **70**(2-3): 141–159. doi:10.1016/j.fishres.2004.08.002.

Maunder, M.N., and Punt, A.E. 2013. A review of integrated analysis in fisheries stock assessment. Fish. Res. **142**: 61–74.

Methot, R.D. 2009. Stock Assessment: Operational Models in Support of Fisheries Management. *In* The Future of Fisheries Science in North America. *Edited by* R.J. Beamish and B.J. Rothschild. Springer Netherlands, Dordrecht. pp. 137–165.

Nielsen, J.R., Kristensen, K., Lewy, P., and Bastardie, F. 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. doi:10.1371/journal.pone.0099151.

Ovaskainen, O., Abrego, N., Halme, P., and Dunson, D. 2016. Using latent variable models to identify large networks of species-to-species associations at different spatial scales. Methods Ecol. Evol. **7**(5): 549–555. doi:10.1111/2041-210X.12501.

Ovaskainen, O., Hottola, J., and Siitonen, J. 2010. Modeling species co-occurrence by multivariate logistic regression generates new hypotheses on fungal interactions. Ecology **91**(9): 2514–2521.

Perry, A.L., Low, P.J., Ellis, J.R., and Reynolds, J.D. 2005. Climate Change and Distribution Shifts in Marine Fishes. Science **308**(5730): 1912–1915. doi:10.1126/science.1111322.

Pinsky, M.L., Worm, B., Fogarty, M.J., Sarmiento, J.L., and Levin, S.A. 2013. Marine taxa track local climate velocities. Science **341**(6151): 1239–1242.

du Preez, C., and Tunnicliffe, V. 2011. Shortspine thornyhead and rockfish (Scorpaenidae) distribution in response to substratum, biogenic structures and trawling. Mar. Ecol. Prog. Ser. **425**: 217–231. doi:10.3354/meps09005.

R Core Team. 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Available from https://www.R-project.org/.

Rooper, C.N., Sigler, M.F., Goddard, P., Malecha, P., Towler, R., Williams, K., Wilborn, R., and Zimmermann, M. 2016. Validation and improvement of species distribution models for structure-forming invertebrates in the eastern Bering Sea with an independent survey. Mar. Ecol. Prog. Ser. **551**: 117–130. doi:10.3354/meps11703.

Shelton, A.O., Thorson, J.T., Ward, E.J., and Feist, B.E. 2014. Spatial semiparametric models improve estimates of species abundance and distribution. Can. J. Fish. Aquat. Sci. **71**(11): 1655–1666. doi:10.1139/cjfas-2013-0508.

Skaug, H., and Fournier, D. 2006. Automatic approximation of the marginal likelihood in non-Gaussian hierarchical models. Comput. Stat. Data Anal. **51**(2): 699–709.

Thorson, J.T., Fonner, R., Haltuch, M., Ono, K., and Winker, H. In press-a. Accounting for spatiotemporal variation and fisher targeting when estimating abundance from multispecies fishery data. Can. J. Fish. Aquat. Sci. doi:10.1139/cjfas-2015-0598.

Thorson, J.T., Ianelli, J.N., Larsen, E., Ries, L., Scheuerell, M.D., Szuwalski, C., and Zipkin, E. In press-b. Joint dynamic species distribution models: a tool for community ordination and spatiotemporal monitoring. Glob. Ecol. Biogeogr.

Thorson, J.T., Pinsky, M.L., and Ward, E.J. 2016. Model-based inference for estimating shifts in species distribution, area occupied and centre of gravity. Methods Ecol. Evol. **7**(8): 990–1002. doi:10.1111/2041-210X.12567.

Thorson, J.T., Scheuerell, M.D., Shelton, A.O., See, K.E., Skaug, H.J., and Kristensen, K. 2015a. Spatial factor analysis: a new tool for estimating joint species distributions and correlations in species range. Methods Ecol. Evol. **6**(6): 627–637. doi:10.1111/2041-210X.12359.

Thorson, J.T., Shelton, A.O., Ward, E.J., and Skaug, H.J. 2015b. Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. ICES J. Mar. Sci. J. Cons. **72**(5): 1297–1310. doi:10.1093/icesjms/fsu243.

Thorson, J.T., Stewart, I., and Punt, A. 2011. Accounting for fish shoals in single- and multi-species survey data using mixture distribution models. Can. J. Fish. Aquat. Sci. **68**(9): 1681–1693.

Thorson, J.T., and Ward, E.J. 2014. Accounting for vessel effects when standardizing catch rates from cooperative surveys. Fish. Res. **155**: 168–176. doi:10.1016/j.fishres.2014.02.036.

Tissot, B.N., Wakefield, W.W., Hixon, M.A., and Clemons, J.E.. 2008. Twenty years of fish-habitat studies on Heceta Bank, Oregon. Mar. Habitat Mapp. Technol. Alsk. Alsk. Sea Grant Coll. Program Univ. Alsk. Fairbanks CD-ROM.

Walters, C. 2003. Folly and fantasy in the analysis of spatial catch rate data. Can. J. Fish. Aquat. Sci. **60**(12): 1433–1436.

Warton, D.I., Blanchet, F.G., O’Hara, R.B., Ovaskainen, O., Taskinen, S., Walker, S.C., and Hui, F.K. 2015. So Many Variables: Joint Modeling in Community Ecology. Trends Ecol. Evol.

Yoklavich, M.M., and O’Connell, V. 2008. Twenty years of research on demersal communities using the Delta submersible in the Northeast Pacific. Mar. Habitat Mapp. Technol. Alsk.: 143–155.

Zipkin, E.F., Andrew Royle, J., Dawson, D.K., and Bates, S. 2010. Multi-species occurrence models to evaluate the effects of conservation and management actions. Biol. Conserv. **143**(2): 479–484. doi:10.1016/j.biocon.2009.11.016.

Table 1 – List of taxa (common and scientific name), the abbreviations used to indicate taxa in plots, and the total number of encounters during 2003-2014.

|  |  |  |  |
| --- | --- | --- | --- |
| Common name | Scientific name | Plotting code | Encounters |
| structure-forming invertebrates (SFI) | - | SFI | 6383 |
| longspine thornyhead | *Sebastolobus altivelis* | L. spine | 2758 |
| shortspine thornyhead | *Sebastolobus alascanus* | S. spine | 3891 |
| darkblotched rockfish | *Sebastes crameri* | Dark | 1338 |
| Pacific Ocean perch | *Sebastes alutus* | POP | 547 |
| sharpchin rockfish | *Sebastes zacentrus* | Sharp | 490 |
| splitnose rockfish | *Sebastes diploproa* | Split | 1619 |
| stripetail rockfish | *Sebastes saxicola* | Stripe | 1630 |
| greenspotted rockfish | *Sebastes chlorostictus* | Green | 434 |

Table 2 – Results from a 10-fold cross-validation experiment comparing single-species models to a multi-species VAST model that was estimated for all species simultaneously, as well as the ratio of predictive probability for the multispecies model relative to the single-species model (a positive value indicates a better fit).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Partition | Number of cross-validation samples | Predictive negative log-likelihood | | Ratio |
|  |  | Single-species VAST model | Multi-species VAST model |  |
| 1 | 6889 | 8459.71 | 8078.84 | 1.057 |
| 2 | 6832 | 8560.26 | 8209.90 | 1.053 |
| 3 | 6835 | 8768.53 | 8377.45 | 1.059 |
| 4 | 6890 | 8203.66 | 7815.19 | 1.058 |
| 5 | 6799 | 8335.34 | 8009.96 | 1.049 |
| 6 | 6828 | 8548.48 | 8154.91 | 1.059 |
| 7 | 6800 | 8426.61 | 8133.02 | 1.044 |
| 8 | 6997 | 8688.89 | 8335.62 | 1.052 |
| 9 | 6743 | 8439.27 | 8077.01 | 1.055 |
| 10 | 6859 | 8594.30 | 8271.42 | 1.048 |

Fig. 1 – Estimated density function (averaged across all years for each taxon) for structure-forming invertebrates, six *Sebastes*, and two *Sebastolobus* species using the VAST model (see Table 1 for plotting code for taxa; inset colorbar shows average log-density in )

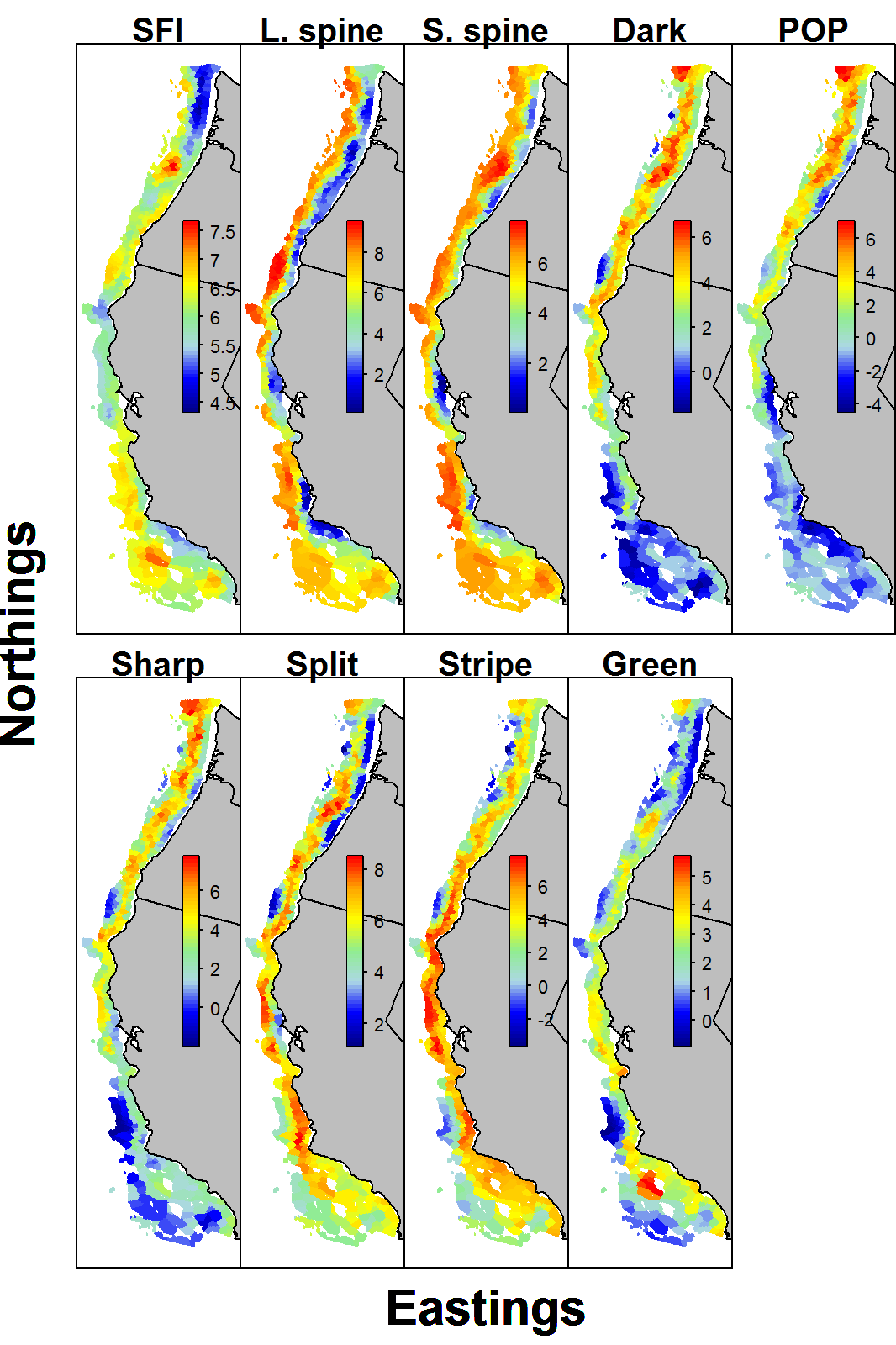


Fig. 2 – Analytic estimate of correlation among fishes and SFI using the VAST model (see Table 1 for taxa abbreviations) for encounter probabilities (,top panel) or positive catch rates (,bottom panel). Numbered columns correspond to the species groups indicated by the row labels, ordered from top to bottom.

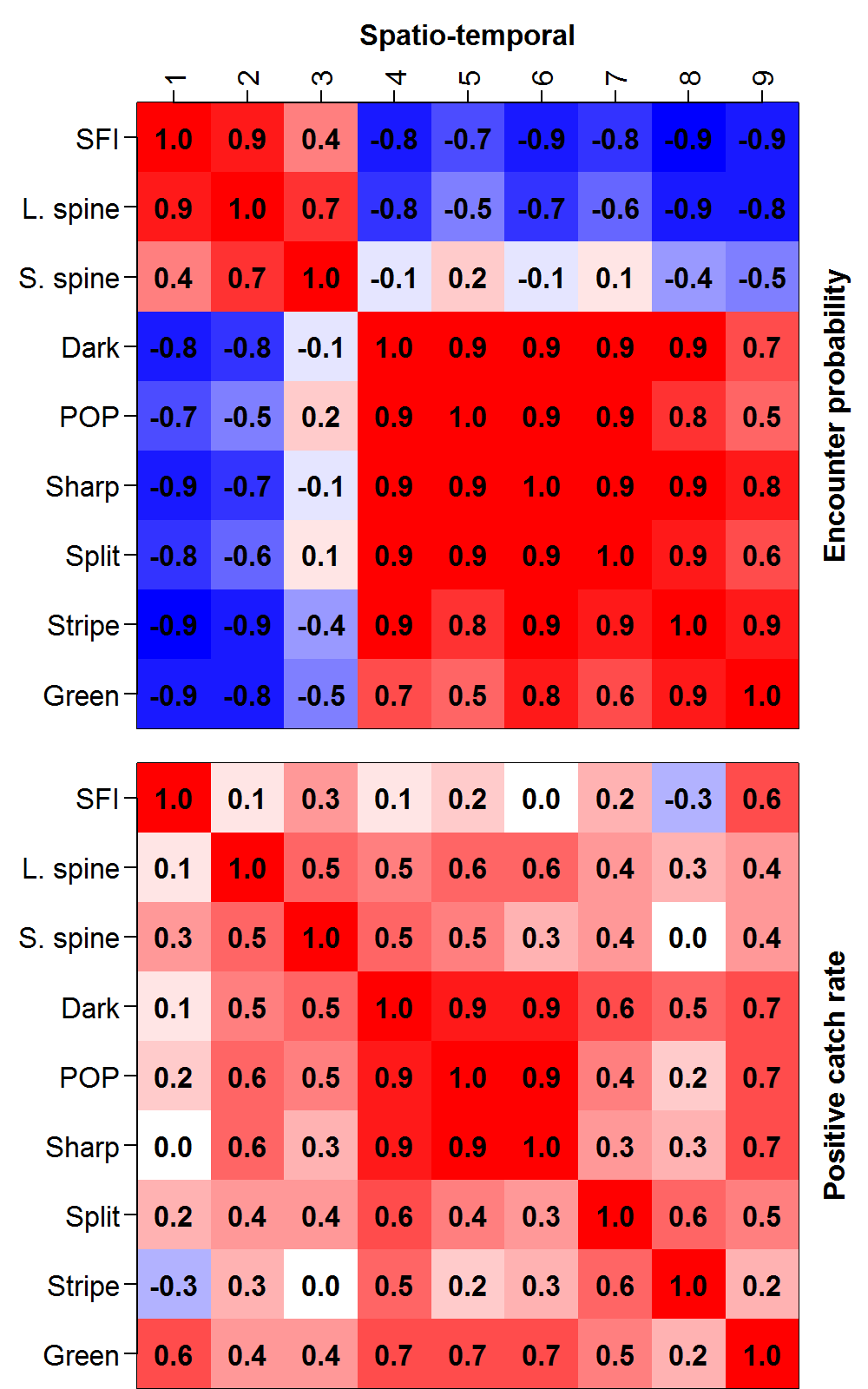


Fig. 3 – Relative log-biomass (in log-metric tonnes) for each species using single-species (grey) or multispecies (red) modelling (note different y-axis range for each species and see Table 1 for taxa abbreviations; top row: SFI and *Sebastolobus* group; middle row: northern *Sebastes* group; bottom row: coastwide *Sebastes* group).

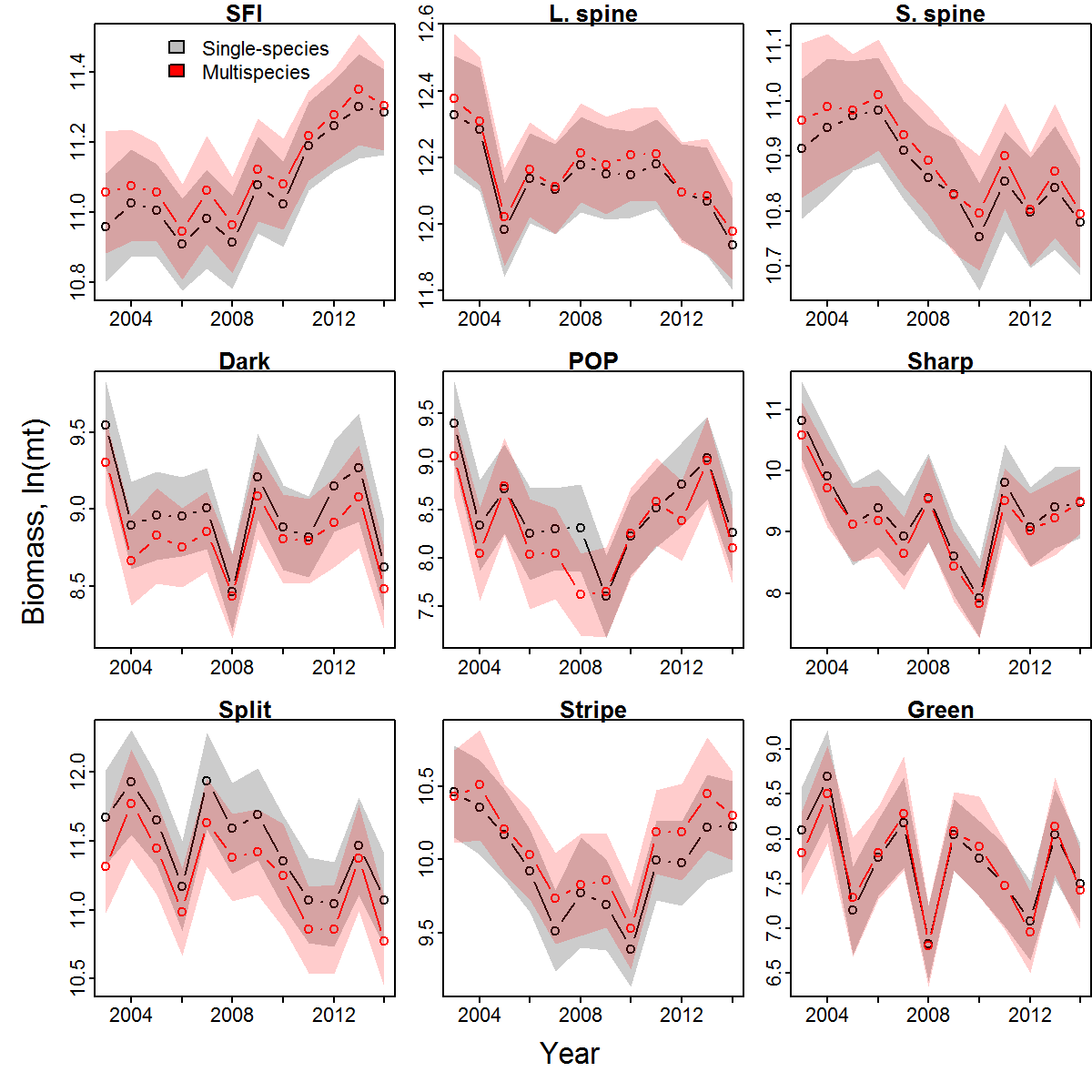


Fig. 4 – Northward center-of-gravity (COG, in kilometres north of equator) for each species using single-species (grey) or multispecies (red) modelling (see Table 1 for taxa abbreviations, and Fig. 3 caption for details).

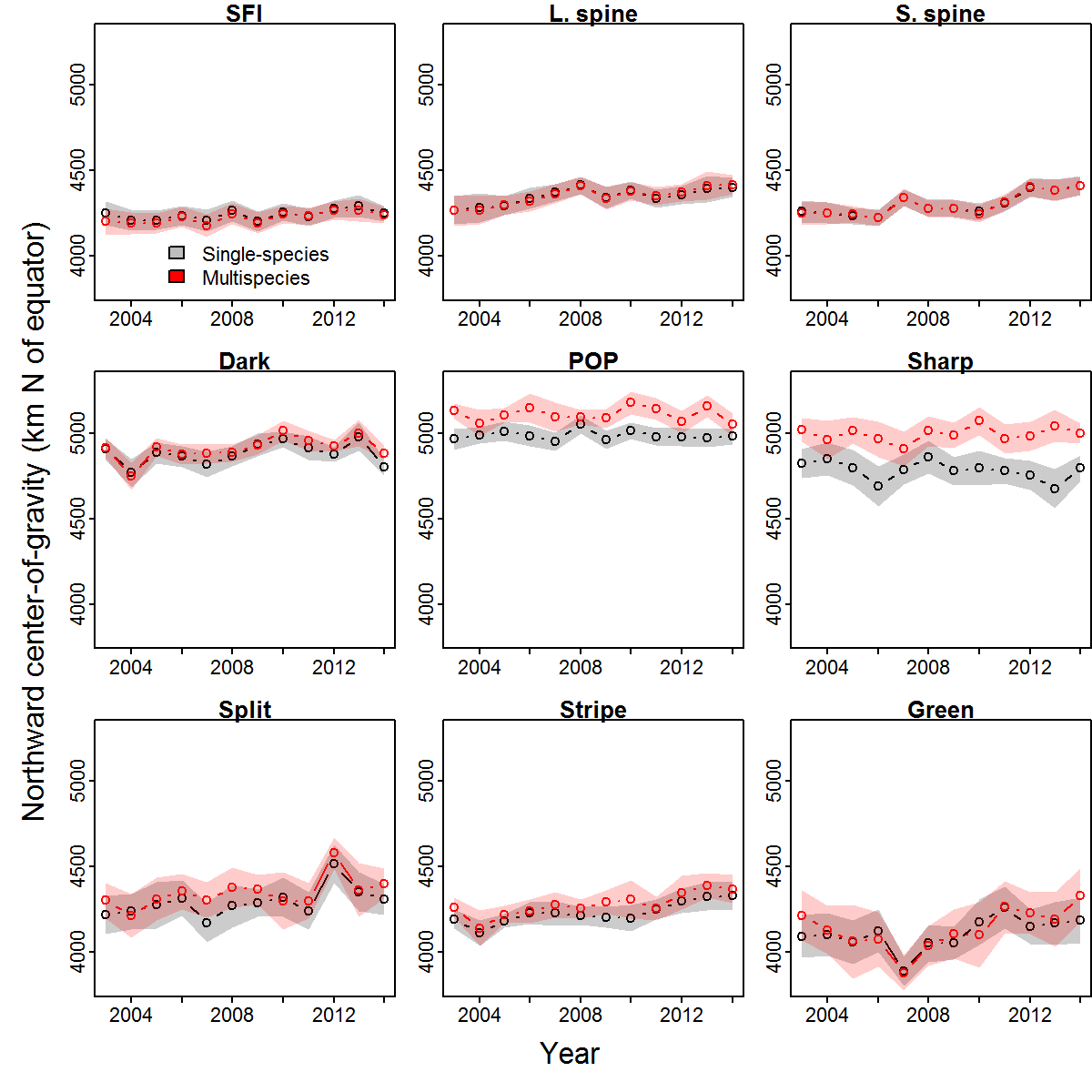


Fig. 5 – Comparison of standard errors for biomass index (top row) and center-of-gravity (COG, bottom row) estimates from single-species and multi-species and models, where the scatterplots compare standard-error estimates for each year and taxon (1st column) and the histograms show the difference between these standard error estimates estimates (2nd column; where a positive value indicates that the multi-species model had a larger standard error than the single-species model for that taxon and year). The dotted line in each histogram indicates the median difference in standard errors between models.

