NRC Research Press
ARTICLE

Accounting for spatiotemporal variation and fisher targeting when estimating abundance from multispecies fishery data[1]

James T. Thorson, Robert Fonner, Melissa A. Haltuch, Kotaro Ono, and Henning Winker

Abstract: Estimating trends in abundance from fishery catch rates is one of the oldest endeavors in fisheries science. However, many jurisdictions do not analyze fishery catch rates due to concerns that these data confound changes in fishing behavior (adjustments in fishing location or gear operation) with trends in abundance. In response, we developed a spatial dynamic factor analysis (SDFA) model that decomposes covariation in multispecies catch rates into components representing spatial variation and fishing behavior. SDFA estimates spatiotemporal variation in fish density for multiple species and accounts for fisher behavior at large spatial scales (i.e., choice of fishing location) while controlling for fisher behavior at fine spatial scales (e.g., daily timing of fishing activity). We first use a multispecies simulation experiment to show that SDFA decreases bias in abundance indices relative to ignoring spatial adjustments and fishing tactics. We then present results for a case study involving petrale sole (*Eopsetta jordani*) in the California Current, for which SDFA estimates initially stable and then increasing abundance for the period 1986–2003, in accordance with fishery-independent survey and stock assessment estimates.

Résumé : L'estimation des tendances d'abondance à partir de taux de prises des pêches est une des pratiques les plus anciennes des sciences halieutiques. De nombreux pays n'analysent cependant pas les taux de prises des pêches parce que ces données pourraient ne pas distinguer les modifications des comportements de pêche (ajustements du lieu de pêche ou de l'utilisation d'engins) des tendances d'abondance. Nous avons donc développé un modèle d'analyse des facteurs de dynamique spatiale (SDFA) qui décompose la covariation des taux de prises multiespèces en différentes composantes représentant la variation spatiale et le comportement de pêche. La SDFA estime les variations spatiotemporelles de densité de poissons pour des espèces multiples et tient compte du comportement de pêche à de grandes échelles spatiales (c.-à-d., le choix du lieu de pêche), tout en contrôlant pour le comportement des pêcheurs à petites échelles spatiales (p. ex. l'heure du jour de l'activité de pêche). Nous utilisons d'abord une expérience de simulation multiespèces pour démontrer que la SDFA réduit le biais dans les indices d'abondance par rapport à la non-prise en compte des ajustements spatiaux et des tactiques de pêche. Nous présentons ensuite les résultats pour une étude de cas portant sur la platija prtrale (*Eopsetta jordani*) dans le courant de Californie, pour laquelle la SDFA estime une abondance initialement stable, puis croissante pour la période 1986–2002, ce qui concorde avec les estimations tirées d'évaluations indépendantes de la pêche et d'évaluations des stocks. [Traduit par la Rédaction]

Introduction

Many developed and developing countries collect detailed information from fishers regarding the location, timing, and catches from fishing operations. This fishing information is typically used to estimate total catch of a given species by a fishery (Somers et al. 2014). Total catch is then often used as a primary input into stock assessment models that are used to estimate population abundance and productivity for informing fisheries management (Methot and Wetzel 2013). Results from stock assessment models are generally used to regulate fishing activities by imposing limits on fishery harvest, or the timing or location of fishing or restrictions on fishing gear.

For over 60 years, fisheries scientists have tried to estimate information about trends in population abundance for exploited fishes from data regarding fishery catches and effort (Beverton and Holt 1957; Smith 2007). Fishery information generally includes the catch per fishing operation (e.g., a fishing trip or trawl tow), and in some cases, changes in catch per effort between years will be informative about changes in population abundance. For example, the common catch equation

$$\mathbb{E}(C) = QFD$$

describes the expected value of catch C as the product of catchability (Q), nominal fishing effort (F), and local density (D). If fishing effort is known and catchability is constant, an increase in catch is

---

Received 22 December 2015. Accepted 3 April 2016.

J.T. Thorson and M.A. Haltuch.* Fisheries Resource Analysis and Monitoring Division, Northwest Fisheries Science Center, National Marine Fisheries Service, NOAA, Seattle, WA, USA.
R. Fonner.* Conservation Biology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, NOAA, Seattle, WA 98112, USA.
K. Ono.* School of Aquatic and Fishery Sciences, Box 355020, University of Washington, Seattle, WA 98195-5020, USA.
H. Winker.* South African National Biodiversity Institute (SANBI), Kirstenbosch Research Centre, Claremont 7735, South Africa; Centre for Statistics in Ecology, Environment and Conservation (SEEC), Department of Statistical Sciences, University of Cape Town, Private Bag X3, Rondebosch, 7701, South Africa.

Corresponding author: James T. Thorson (email: James.Thorson@noaa.gov).

*RF, MAH, KO, and HW contributed equally to this contribution and are listed alphabetically.

[1] This article is being published as part of the special issue "Space Oddity: Recent Advances Incorporating Spatial Processes in the Fishery Stock Assessment and Management Interface" arising from a related theme session at the 145th Annual Meeting of the American Fisheries Society, Portland, Oregon, USA, August 2015.

Can. J. Fish. Aquat. Sci. **74**: 1794–1807 (2017) dx.doi.org/10.1139/cjfas-2015-0598

associated with increasing fish abundance. Therefore, a regression model can be fitted to records of catch data, and the resulting estimates of density can be used as an index of abundance in a stock assessment model. Data regarding fishing characteristics that affect fishing effort or catchability (e.g., vessel identity, the timing of fishing, etc.) can be included as covariates in the regression model (Maunder and Punt 2004). If this is done, variation caused by these covariates is "controlled for" (i.e., filtered out) when estimating the resulting index of abundance. This process is typically referred to as "standardizing" the index of abundance (Beverton and Holt 1957; Maunder and Punt 2004).

However, standardizing catch and effort data for a fishery poses many complications. Fishers will often restrict fishing to areas with high fish density, and if the fish become more concentrated over time, this will generally result in elevated catch rates and mask evidence of declining population abundance (Walters and Maguire 1996). Analysts may partially account for this by standardizing fishery data using a spatial index standardization method (Walters 2003; Carruthers et al. 2011). Furthermore, the common catch equation is a simplification of the Cobbs–Douglas production function, given the assumption that fishing effort and density both have a linear effect (Wilberg et al. 2010). Catches may, in fact, be nonlinearly related to local densities and fishing effort (Harley et al. 2001), although these nonlinear effects are not generally estimable without auxiliary information (Kotwicki et al. 2013). Finally, catchability may vary among fishing operations in ways that are not directly observable, e.g., due to small changes in timing of fishing operations.

The catchability of fishing operations may also vary due to a number of technical and behavioral characteristics of fishing (Pelletier and Ferraris 2000; Wilberg et al. 2010; Abbott et al. 2015). Research suggests that fishers can affect the assortment of species caught in a multispecies fishery by modifying the location, timing, and gear characteristics of their fishing activity (Sanchirico et al. 2006; Abbott et al. 2015). Fishers make discrete operational choices that influence how the catch is distributed across space, time, and species. Before a given fishing season begins, eligible fishers choose whether to participate in the fishery and make vessel capitalization decisions (e.g., repairs, equipment purchases). Once the season starts, fishers must then decide where to fish. The expected profit framework for explaining fisher location decisions (Haynie and Layton 2010) assumes that fishers associate each alternative fishing location with an expected level of profit based on the expected catch (i.e., catch rate and species composition) and expected costs of production (e.g., steaming distance, labor, risk to gear) in that area, and that fisher locational choices are made to maximize expected profits. Upon harvesting fish at their initial location, fishers compare the expected catch with what they are catching. Based on the realization of expectations, fishers can adjust their operation through coarse- or fine-scale spatial adjustments, timing adjustments (e.g., soak time, time of day), or gear adjustment or by returning to port. The incentives motivating fisher choices are governed by factors across economic (e.g., input prices, ex-vessel fish prices), environmental (e.g., expected distribution of fish, steaming distances, weather), and regulatory (e.g., spatial closures, catch and effort limits, gear restrictions) dimensions. Understanding how these factors influence fisher behavior in a given fishery can inform the development and interpretation of spatiotemporal models of catch distributions.

While fishers generally prefer to harvest species that yield high profits per unit effort, perfect targeting (the ability of fishers to select the species they catch) is often not possible in multispecies fisheries. In other words, these fisheries exhibit joint production technology in which fishing effort always yields more than one species. With joint production, fishers target species assemblages that maximize their profit rather than targeting a single species[2]. At large spatial scales, targeting is accomplished though selection of fishing grounds. At fine spatial scales, targeting is accomplished through implementing fishing tactics (e.g., gear, time of day, bearing, small-scale movements). For clarity, this paper refers to coarse spatial scale fisher location choices as "spatial adjustments" and fine spatial scale behavior as "tactics". Multispecies fisheries with joint production technology are associated with spatial clustering of fishers into groups, called métiers[3], which tend to target similar fishing grounds and employ similar tactics.

Information about fishing tactics is increasingly used when standardizing fishery data to estimate an index of abundance (He et al. 1997; Winker et al. 2013). Stephens and MacCall (2004) developed a method that first fits a logistic regression model to the presence or absence of the target species, predicted via catch of nontarget species. Fishing operations with a low probability of catching the target species are then excluded in a second-stage model that estimates abundance for the target species. Other authors have applied ordination techniques to multispecies catch rate information in a first-stage model (Pelletier and Ferraris 2000). Results from this ordination are then used as either a categorical (Maynou et al. 2003) or continuous (Winker et al. 2013) variable in a second-stage regression model. These two-stage modelling techniques have several drawbacks, including

1. an inability to propagate uncertainty from the model defining fishing tactics into the subsequent regression model used to estimate abundance;
2. a lack of transparency regarding model design, as well as an inability to simultaneously visualize model residuals from both the first- and second-stage models;
3. difficulties in determining whether the variable or factor representing fishing tactics is orthogonal to variables of interest in the second model or, if not orthogonal, what effect this lack of independence has on model performance when estimating abundance;
4. the fact that many subjective decisions must be made, i.e., regarding thresholds for excluding tows (Stephens and MacCall 2004), transforming data when estimating fishing tactics (Deporte et al. 2012; Winker et al. 2013), or excluding observations with zero catches, and the effects of these subjective decisions are difficult to evaluate.

Despite these drawbacks, research suggests that accounting for fishing tactics leads to performance improvements when estimating abundance indices from fishery data (Haltuch et al. 2013; Winker et al. 2014).

In this study, we seek to develop a theoretically interpretable and statistically robust method for simultaneously estimating fishing tactics and relative fish abundance when standardizing fishery-dependent catch rate data. The proposed "spatial dynamic factor analysis" (SDFA) model estimates spatial and temporal variation in abundance for multiple fishes caught in a multispecies fishery, while simultaneously accounting for residual correlations in fishery catch rates caused by fishing behavior. After reviewing the context and underlying motivations driving fisher behavior, we use theoretical arguments and a simulation experiment to show that the SDFA model accounts for fisher behavior occurring at coarse spatial scales while controlling for fishing tactics at fine spatial scales (Table 1). We then provide a case-study example involving the winter fishery for petrale sole (*Eopsetta jordani*) 1986–2003, in which the resulting fishery-dependent abundance index shows temporal patterns that

---

[2] In some cases, the profit-maximizing assemblage may be the one with the highest overall density of the most valuable species, though this is not always the case. For example, it may be beneficial for fishers to avoid a protected species.

[3] Pascoe (2015) defines "métier" as "a fishing activity or strategy defined by area and gear, and therefore associated with a particular catch composition as a result".

Table 1. A summary of the influence of fisher behavior on catch distributions at large and small spatial scales.

	Mechanisms	Model treatment
Spatial adjustments (i.e. the allocation of effort among sites)	• Initial location choice based on expected profit • Spatiotemporal adjustments in fishing location related to changes in relative ex-vessel prices of species, input costs, and regulations over time • Changes in fishing location due to new information obtained from prior fishing (e.g., avoiding areas with low catch rates)	$\mathrm{Cov}(\log(D_p)) = \mathbf{AA}^{\mathrm{T}}$
Tactics (i.e. residual correlations in catch rates at a given site)	• Fine-scale spatial adjustments to seek a more favorable species composition and higher catch rates once catch is observed at initial location • Changes in timing of fishing activity (e.g., daytime, nighttime, crepuscular) • Changes in fishing operations, e.g., bearing and speed • Changes in fishing gear (e.g., bait type, hook type, mesh size)	$\mathrm{Cov}(\log(F_p)) = \mathbf{BB}^{\mathrm{T}}$

are consistent with survey data and stock assessment estimates. This case study identifies that fine-scale targeting in this fishery generates a trade-off between petrale sole and dover sole (*Microstomus pacificus*), while the fishery can also adjust the spatial distribution of fishing to target locations with high petrale density. We provide an R package, Spatial_DFA, for implementing SDFA for other cases (https://github.com/James-Thorson/spatial_DFA) and use release number 1.0.0 in the following analysis.

Methods

Study goals

We seek to estimate an index that is proportional to abundance for a given species using records of catch and fishing effort for multiple species that are caught within a fishery. To do this, we seek to account for four processes that contribute to patterns of fishery catch in multispecies fisheries:

1. spatial variation in density for multiple species: different species will typically occupy different habitats when averaged over long time periods; the spatial mosaic of occupied habitats arises from spatial variation in habitat quality, dispersal ability, and competitive interactions (Soberón 2007);
2. autocorrelated, spatiotemporal variation in density for multiple species: a species' distribution will not generally match its mean distribution in any given year, and changes in spatial distribution among years can be driven by changes in biological (e.g., spawning dynamics), physical (e.g., ocean temperature), or human-caused (e.g., fishing) processes; deviations away from a species' long-term spatial distribution may persist over several sequential years, for example, if juveniles and adults of a species have different spatial distributions and the species has high variability in cohort strength, then a strong cohort will likely cause autocorrelation in spatiotemporal variation;
3. correlations in spatiotemporal variation among species: both spatial and spatiotemporal variation will, in some cases, be correlated among species; these correlations arise whenever species have similar habitat associations (Kissling et al. 2012);
4. correlated catch rates arising from fisher tactics: fishers using trawl or hook-and-line gears sometimes seek to catch one or more high-value species, but also incidentally catch other species due to the joint production nature of many multispecies fisheries; research suggests that fishers can modify the assortment of species caught during a given occasion by moving to a different fishing location or by modifying fishing tactics within a given fishing location (Sanchirico et al. 2006; Abbott et al. 2015). The existence of fine-scale fishery tactics implies that there will be multispecies correlations in catch rates even after accounting for spatiotemporal variation in density for each species. Negative residual correlations may also arise due to competition among species for fishing hooks.

In the following, we distinguish between spatiotemporal and spatially stratified models.

- A spatiotemporal model estimates the magnitude and impact of autocorrelation between nearby sites and nearby time periods. As the number of locations increases, the model estimates a smooth function representing spatial distribution, as well as changes in spatial distribution over time (Thorson et al. 2015*a*).
- A spatially stratified model estimates mean density within different predefined areas. As the number of strata increases, the model will eventually fail to converge as some strata have no data to inform estimates of density (Thorson and Ward 2013).

We seek to estimate the distance over which density is correlated (corresponding to the size of habitat patches) and therefore seek to account for the preceding four processes using a spatiotemporal model.

Model structure

We use a statistical model that accounts for fishing behavior, spatial and spatiotemporal variation, and correlations among species, called "spatial dynamic factor analysis" (SDFA). SDFA explains density of P species as a log-linear combination of J factors ($0 \leq J \leq P$) in which each factor represents unobserved spatial variation in density for a species assemblage. The jth factor follows an autoregressive spatiotemporal process:

$$\psi_j(s, t+1) = \rho\psi_j(s, t) + \omega_j(s) + \xi_j(s, t)$$

where $\psi_j(s, t)$ is the value of factor j at site s and time t, ρ is the magnitude of autocorrelation, $\omega_j(s)$ is the mean value of factor j at site s, and $\xi_j(s, t)$ represents spatiotemporal variation in that factor (see Table 2 for a list of all symbols used in this model). Both spatial and spatiotemporal variations follow a spatially correlated stochastic process:

$$\omega_j \sim \mathrm{GRF}(0, R(|\mathbf{H}\Delta s|;\, \kappa))$$

$$\xi_j(t) \sim \mathrm{GRF}(0, R(|\mathbf{H}\Delta s|;\, \kappa))$$

where GRF(0, R) is a zero-mean Gaussian random field with correlation function R (see Thorson et al. (2015*a*) for more details regarding Gaussian random fields), R is a Matérn correlation function given distance Δs between two locations and geometric anisotropy $\mathbf{H}$, where parameter κ governs the distance over which covariance drops to 10% of its pointwise value (Thorson et al. 2015*c*).

Log-expected density, $\log(d_p(s, t))$, for species p is then a linear combination of factors and measured spatiotemporal covariates:

$$\log(d_p(s, t)) = \sum_{j=1}^{J} A_{p,j}\psi_j(s, t) + \sum_{l=1}^{L} \gamma_{p,l}(t)x_l(s, t)$$

where $\mathbf{A}$ is a $P \times J$ matrix representing the impact $A_{p,j}$ of factor j on density for species p (termed the "loadings matrix"), $x_l(s, t)$ is the

Table 2. List of theoretical variables, parameters, data, and indices used in the spatial dynamic factor analysis model.

Name	Symbol
Theoretical variables	
Catch	C
Fishery catchability	Q
Fishing effort	F
Fish density	D
Model parameters	
Spatial dynamic factor	$\psi_j(s, t)$
Spatial variation	$\omega_j(s)$
Spatiotemporal variation	$\xi_j(s, t)$
Temporal autocorrelation for each factor	ρ
Matérn correlation function	R
Geometric anisotropy	$\mathbf{H}$
Geostatistical range	κ
Log-expected density	$d_p(s, t)$
Loadings matrix for density covariation	$\mathbf{A}$
Coefficients representing linear effect of density covariates on log-density	$\gamma_{p,l}(t)$
Loadings matrix for targeting covariation	$\mathbf{B}$
Factor representing targeting	$\varepsilon_k(i)$
Log-expected catch rates	$\lambda_p(i)$
Coefficient of variation for positive catch rates	$\upsilon_{p,1}$
Expected rate of decrease of zero catches with increasing density	$\upsilon_{p,2}$
Index of abundance	$I_p(t)$
Data	
Density covariate	$x_l(s, t)$
Catchability covariate	$y_m(i)$
Observed catch rates	$c_p(i)$
Area associated with each location	$a(s)$
Dimensions	
Number of spatial dynamic factors	J
Number of locations	S
Number of time intervals	T
Number of species	P
Number of density covariates	L
Number of targeting factors	K
Number of catchability covariates	M
Indices	
Spatial dynamic factor	j
Location	s
Time interval	t
Species number	p
Density covariate	l
Targeting factor	k
Catchability covariate	m

lth measured variable (i.e., an intercept, or water temperature) at site s and time t, and $\gamma_{p,l}(t)$ is the effect of covariate l on density for species p in year t. Loadings matrix $\mathbf{A}$ governs the covariation in log-density among species and therefore ensures that density estimates for each species are informed by catch rates for every other species. We note that the variance of both spatial ω_j and spatiotemporal $\xi_j(t)$ random effects is one, such that the variance of spatial $\mathbf{A}\boldsymbol{\omega}_j$ and spatiotemporal $\mathbf{A}\boldsymbol{\xi}_j$ variation is controlled by the value of $\mathbf{A}$. Log-expected catch, $\log(\lambda_p(i))$, for species p and fishing occasion i is then a linear combination of local log-density, variation in catchability caused by fishing tactics, and observed variables associated with variation in catchability (e.g., vessel ID):

$$\log(\lambda_p(i)) = \log(d_p(s_i, t_i)) + \sum_{k=1}^{K} B_{p,k}\varepsilon_k(i) + \sum_{m=1}^{M} \nu_{p,m} y_m(i)$$

where $y_m(i)$ is the mth measured variable affecting catchability for fishing occasion i, $\nu_{p,m}$ is the effect of covariate m on catch rates for species p, and $\mathbf{B}$ is a $P \times K$ loadings matrix representing the impact $B_{p,k}$ on species p of small-scale variables $\varepsilon_k(i)$:

$$\varepsilon_k(i) \sim N(0, 1)$$

such that $\varepsilon_k(i)$ is one or more latent variables representing variation in fishing tactics, where $0 \leq K \leq P$. Variable $y_m(i)$ therefore accounts for measured processes affecting catchability (e.g., fishing vessel size), while $\mathbf{B}$ and $\varepsilon_k(i)$ account for residual, unmeasured variation in catchability. The component of this model representing multispecies spatiotemporal dynamics (i.e., $d_p(s, t)$) has previously been explored by Thorson et al. (2016*a*). The component representing fishing tactics (i.e., $B_{p,k}\varepsilon_k(i)$) has not previously been included in SDFA, but resembles correlations used in other spatiotemporal models (e.g., Kristensen et al. 2014).

Finally, we specify a distribution for residual variation in catch rates:

$$\Pr(C = c_p(i)) = \begin{cases} \exp(-\upsilon_{p,2}\lambda_p(i)) & \text{if } C = 0 \\ (1 - \exp(-\upsilon_{p,2}\lambda_p(i)))\text{Gamma}\left(C; \upsilon_{p,1}^{-2}, \dfrac{\lambda_p(i)}{1 - \exp(-\upsilon_{p,2}\lambda_p(i))}\upsilon_{p,1}^{2}\right) & \text{if } C > 0 \end{cases}$$

where Gamma(C; x, y) is the PDF of a gamma distribution with shape x and scale y, $\upsilon_{p,1}$ is the coefficient of variation for positive catches for species p, and $\upsilon_{p,2}$ controls the relationship between the probability of encountering zero individuals and predicted catch for species p such that the probability of not encountering a species ($C = 0$) is identical to a Poisson distribution with intensity $\upsilon_{p,2}\lambda_p(i)$.

We estimate parameters using conventional techniques for mixed-effect models (Thorson and Minto 2015). Specifically, fixed effects are estimated by identifying their values that maximize the likelihood function, which specifies the probability that data arose given those values for fixed effects when integrating across random effects. We estimate fixed effects using Template Model Builder (TMB; (Kristensen 2014)) from within the R statistical environment (R Core Team 2013). Parameter estimation using TMB involves the following steps: (*i*) we specify a template file that specifies the joint probability, calculated as the product of the probability that the data arose given specified values for fixed and random effects and the probability that random effects arose given values for fixed effects; (*ii*) for input values of fixed effects, TMB optimizes the joint likelihood with respect to random effects; (*iii*) TMB calculates the second derivatives of the joint likelihood with respect to random effects (the Hessian matrix); (*iv*) TMB uses the Hessian matrix to calculate the Laplace approximation to the marginal likelihood function and the gradients of this function with respect to fixed effects; and (*v*) we apply a conventional nonlinear optimizer, using the marginal likelihood function and its gradients, to identify maximum likelihood estimates (MLE) for fixed effects. TMB then predicts random effects at the values that maximize the joint likelihood conditional on the MLE for fixed effects and uses the Hessian of the marginal likelihood with respect to fixed effects and the generalized delta method to calculate standard errors for all fixed and random effects (Kass and Steffey 1989). For computational convenience, we use the stochastic partial differential equation (SPDE) approximation (Lindgren et al. 2011) when calculating the probability of Gaussian random fields (ω_j and $\xi_j(t)$; see Thorson et al. (2014) for more details). Code for implementing the model is available in the Spatial_DFA R package.

Model interpretation

SDFA can be interpreted as a way to factor the catch equation for species p, $\mathbb{E}(C_p) = Q_p F_p D_p$, such that

$$\log(D_p) = \log(d_p(s,t)) = \sum_{j=1}^{J} A_{p,j}\psi_j(s,t) + \sum_{l=1}^{L} \gamma_{p,l}(t)x_l(s,t)$$

$$\log(F_p) = \sum_{k=1}^{K} B_{p,k}\varepsilon_k(i)$$

$$\log(Q_p) = \sum_{l=1}^{M} \nu_{p,m}y_m(i)$$

where F_p is nominal fishing effort after accounting for small-scale fisher targeting, and Q_p is catchability and where residual variation follows the zero-inflated Gamma distribution. The covariance of spatiotemporal variation in log-density among species can then be calculated as

$$\mathrm{Cov}(\log(D_p)) = \mathbf{AA}^{\mathrm{T}}$$

while the covariance of small-scale variation in fishing effort via fisher tactics is

$$\mathrm{Cov}(\log(F_p)) = \mathbf{BB}^{\mathrm{T}}$$

It is impossible to estimate spatiotemporal covariation ($\mathbf{AA}^{\mathrm{T}}$) occurring at a finer spatial scale than the mean minimum distance between sites, so remaining covariation is attributed to fishing tactics ($\mathbf{BB}^{\mathrm{T}}$). The superscript "T" means "transpose. We refer to distances greater than this minimum distance as "coarse spatial scales", and smaller distances as "fine spatial scales".

The model also implies that there is no covariation between density and nominal fishing effort:

$$\mathrm{Cov}(\log(D_p), \log(F_p)) = 0$$

i.e., fishers make a decision about the location of fishing and then make a statistically independent decision about fishing tactics at that location. We note that this property is likely violated in real-world applications, where fishers may make different decisions about fishing tactics depending on the location of fishing. However, a simple model estimating a linear linkage between density and fishing effort, $\mathrm{Cov}(\log(D_p), \log(F_p)) = \alpha$, would result in complete confounding of the link parameter α. We view it as a strength of the SDFA model that it allows explicit consideration of potential mechanistic linkages between density and small-scale fishing tactics, but we leave further exploration as a topic for future research. We hypothesize that the SDFA model will perform well at estimating trends in abundance when changes in species composition arise primarily from the spatial allocation of fishing effort.

Additionally, an index $I_p(t)$ of total abundance for each species p in time t can be calculated:

$$I_p(t) = \sum_{s=1}^{S} a(s)d_p(s,t)$$

where $a(s)$ is the area associated with each site. The standard error of this estimate can then be obtained via conventional asymptotic or likelihood-profile techniques.

Simulation experiment

We first conduct a simulation experiment to explore the performance of the SDFA model. To do so, we generate artificial data sets using a simulation model previously developed by Ono et al. (2015). However, we use additional features that were not previously described in that paper. Specifically, we simulate density for four species within each of 100 cells in a 10 cell × 10 cell grid. Dynamics for each species follows logistic population growth (given cell-specific density dependence), with movement among cells controlled by the species' habitat preference and mobility and where fishing effort and resulting catches are governed by expected revenue given densities and prices for each species in each year. A single habitat feature (e.g., depth) is simulated as a Gaussian random field, and this habitat variable affects each species' density according to species-specific habitat preferences. For each year in a 30-year simulation interval, the simulation model then involves the following steps:

1. simulate the population dynamics of each species using the Schaefer model (Schaefer 1954), where surplus production in each cell is a function of biomass in that cell (i.e., density dependence is local to each cell);
2. calculate the total fishing effort in that year and then distribute fishing effort to each cell given the expected revenue at the start of the year;
3. calculate the corresponding catch in each cell and record catch for use in subsequent analysis while incorporating additional measurement error;
4. decrease population abundance in each cell by the amount of fishery catch;
5. redistribute species abundance among cells based on species' movement preferences.

Further details are provided in Appendices A and B.

We present results for a scenario in which fish prices change over time, spurring changes in species targeting and therefore the spatial distribution of fishing effort. Specifically, fish prices are identical for all four species for years 1–10, but in years 11–20, prices go up for species 1 (price is doubled) and down for species 3 and 4 (price is halved), while remaining constant for species 2. Then, in years 21–30, prices return to their original levels. These changes in price induce changes in the spatial distribution of fishing effort. This simulator is available using the function Generate_scenario_data in the Spatial_DFA R package.

This simulation model generates 30 years of catch and effort data distributed among 100 grid cells. The data frame containing catch and fishing effort in each cell and year is then provided to the estimation model, and estimates of total abundance for each species are compared with the "true" simulated abundance for each species. For each simulated data set, we estimate abundance using four configurations of the SDFA model:

1. default model: excluding both spatial variation and small-scale tactics ($\mathbf{A} = \mathbf{B} = \mathbf{0}$ such that $J = K = 0$);
2. tactics model: estimating only covariation caused by fishing tactics ($\mathbf{A} = \mathbf{0}$ but estimating $\mathbf{B}$ such that $J = 0$ and $K = 4$);
3. spatial model: estimating only spatial variation ($\mathbf{B} = \mathbf{0}$ but estimating $\mathbf{A}$ such that $J = 4$ and $K = 0$); or
4. spatial and tactics model: estimating both spatial variation and fisher tactics (estimating both $\mathbf{A}$ and $\mathbf{B}$ such that $J = K = 4$).

For computational reasons, we do not include model selection for number of factors (J) or tactics variables (K) between 0 or 4, but instead select among major hypotheses (i.e., including or excluding spatiotemporal variation or fisher tactics). We also do not include any catchability covariates (i.e., $M = 0$), and the only density covariate is an intercept for each species and year (i.e., $L = 1$, where $x_l(s,t) = 1$ and $\gamma_{p,l}(t)$ involves estimating $P \times T$ intercept parameters). We recommend that future research further explores the impact of model selection on SDFA performance.

We replicate this simulation experiment 100 times to explore the estimation performance of these four alternative models. We evaluate performance by calculating log of the relative error (LRE) for each model:

$$\mathrm{LRE}_{p,t} = \log\left(\frac{\hat{I}_{p,t}\Big/\sum_{t=1}^{T}\hat{I}_{p,t}}{I_{p,t}\Big/\sum_{t=1}^{T}I_{p,t}}\right)$$

where $\mathrm{LRE}_{p,t}$ is the log of the relative error for species p in year t, $\hat{I}_{p,t}$ and $I_{p,t}$ are the estimated and true abundances, respectively, for that species and year, $T = 30$ is the total number of years such that $\hat{I}_{p,t}\big/\sum_{t=1}^{T}\hat{I}_{p,t}$ and $I_{p,t}\big/\sum_{t=1}^{T}I_{p,t}$ are, respectively, the estimated and true indices of abundance after rescaling each to have a mean of one (i.e., given their treatment as relative indices of abundance).

Case study: petrale sole

We also provide a worked example by using the SDFA model to analyze fishery catch rate data for the winter (November–February) fishery for petrale sole in waters off Oregon and Washington from 1986–2003. We restrict analysis to data for fishery logbook records of catches of four categories (petrale sole, dover sole, sablefish (*Anoplopoma fimbria*), and thornyheads, where the latter comprises catches of both shortspine (*Sebastolobus alascanus*) and longspine (*S. altivelis*) thornyheads), and hypothesize that fishers target either petrale or dover sole, while potentially avoiding sablefish (n = 47 652 catch rate records for each of the four categories). We again do not include any catchability covariates (i.e., $M = 0$) and include an intercept for log-density for each species and year (i.e., $L = 1$, where $x_l(s, t) = 1$ and $\gamma_{p,l}(t)$ involves estimating 4 × 18 parameters). To increase computation speed, we estimate abundance at 400 "knots" (i.e., $S = 400$). To do this, we apply a k-means clustering algorithm to the location of all samples and then assume that density for each sample is identical to estimated density for the nearest knot (following Thorson et al. (2015*c*)). Each knot is associated with a 102.1 km^2 area, on average, so processes occurring at smaller scales are attributed to fishing tactics ("fine scale") and those occurring at larger scales are attributed to spatial adjustments ("coarse scale"). We specify full rank for the species-tactics matrix (i.e., $K = 4$) and use the Akaike information criterion to select the number of spatial factors (i.e., selection among $J = \{0, 1, 2, 3, 4\}$). We then extract the estimate of total abundance for petrale sole from the selected model.

Next, we compare estimates of total abundance for the selected model with estimates of total abundance from the triennial bottom-trawl groundfish survey. To do so, we fit a previously developed geostatistical index standardization model (Thorson et al. 2015*c*) to survey data collected during June to September from years 1980, 1983, 1986, 1989, 1992, 1995, 1998, and 2001 collected by the Alaska Fisheries Science Center (Weinberg 2002). This index standardization model separately analyzes the proportion of survey tows that encounter petrale and the density of petrale when encountered using a delta generalized linear mixed model (delta-GLMM) design. The seasonal timing of the triennial survey changed between 1992 and 1995, so we include calendar date (standardized to have zero mean and unit variance) as a linear covariate in the submodels for encounter probability and density given encounters.

We also compare the estimates of petrale abundance from SDFA analysis of fishery catch rate data with estimates of petrale abundance from the latest stock assessment in 2013 (Haltuch et al. 2013). However, we modify the assessment to exclude all fishery-dependent catch-per-unit-effort (CPUE) information, so that estimates from the stock assessment are statistically independent of the fishery-dependent catch rate data. This stock assessment synthesizes many other data sources, including the triennial survey, as well as age- and length-composition samples, to estimate abundance for petrale sole in the California Current.

Finally, we conduct a 10-fold cross-validation experiment to evaluate performance of the SDFA model. Cross-validation involves dividing the entire data set of fishery catch rates into 10 equal-sized partitions. For the first replicate, we exclude data for petrale sole from partition 1 (called the "testing set"), while fitting to data for partitions 2–10 for petrale and all partitions (1–10) for all other species (called the "training set"). This fitted model is used to calculate the natural logarithm of probability of data for petrale in the testing set. We refer to the log-probability of the testing set, divided by the number of observations in the testing set, as the "predictive score". This process is then repeated for all 10 partitions using the four alternative model configurations (default, spatial, tactics, and spatial and tactics models). The model that has the highest predictive score is inferred to be the most parsimonious model for out-of-sample prediction (Gelman et al. 2013).

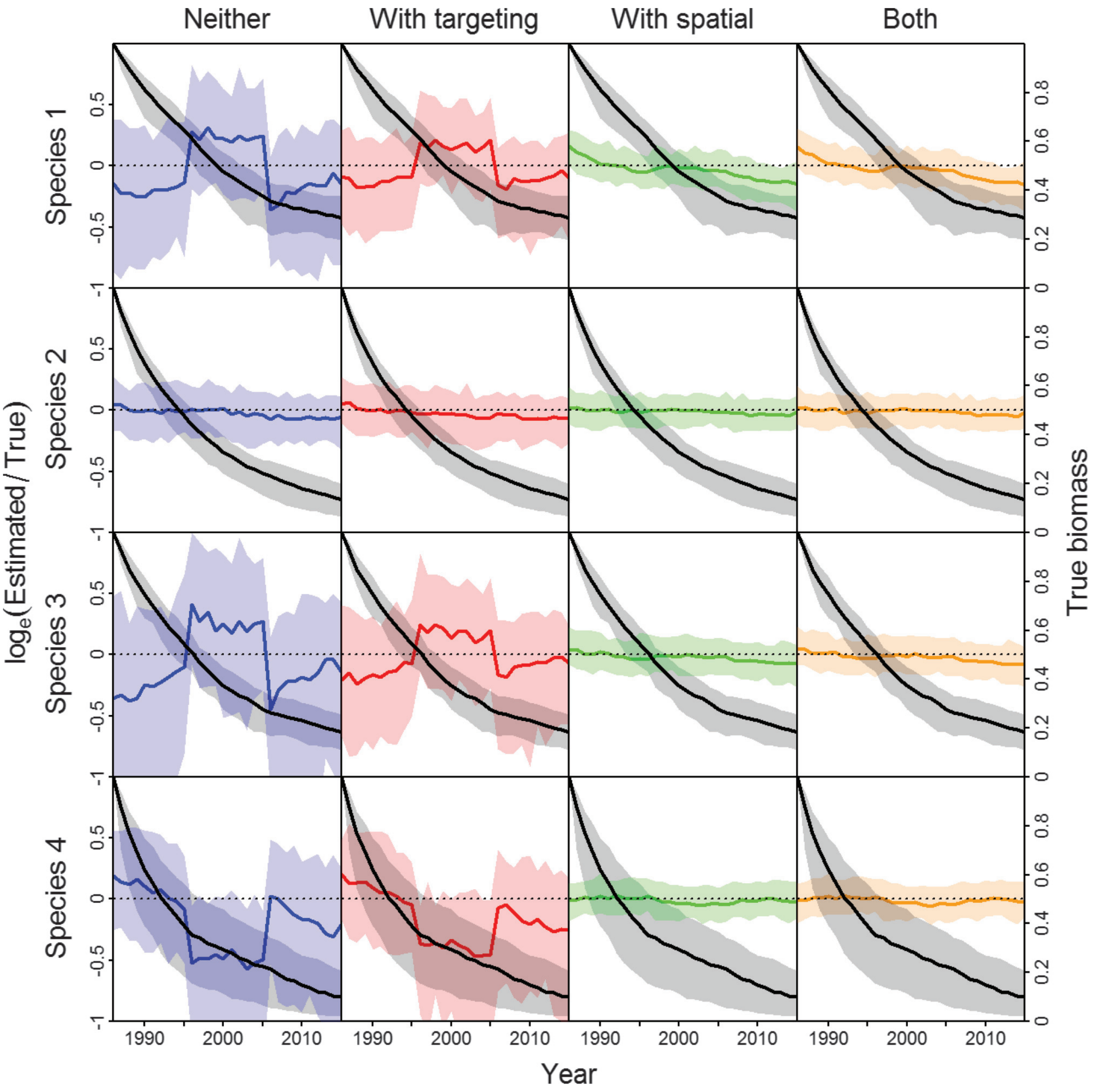
Results

Examining results from the simulation experiment (Fig. 1) shows that all models are approximately unbiased for species 2, which has no change in pricing among years. However, the model without spatial variation or tactics (Fig. 1, 1st (left) column) is highly biased for the species that have changes in tactics (species 1, 3, and 4). This bias is positive for species 1 and 3 in years 11–20 and negative for years 1–10 and 21–30, where changes in bias coincide with changes in the spatial allocation of fishing (caused by changes in relative price). This pattern of bias is reversed for species 4. The model with only tactics (Fig. 1, 2nd column) somewhat decreases bias for species 1, 3, and 4 but is still substantially biased. The models with spatial variation or both spatial and tactics (Fig. 1, 3rd and 4th columns) are least biased but still show a small positive bias in early years and negative bias in late years for species 1, and results show that adding tactics (i.e., 4th column) offers little improvement over the model with just spatial variation (3rd column). A follow-up analysis (Appendix C) shows that models with spatial variation or both spatial and tactics have decreased performance when using data that are spatially aggregated prior to analysis.

Model selection indicates that the model with four dynamic factors ($J = 4$) has the most parsimonious fit for the winter petrale fishery. Mapping density for the fishery-dependent index shows that density is highly patchy, with elevated density in small areas off northern and southern Oregon, and that the hotspot in southern Oregon shows a particularly large increase in abundance from 1993 onward (Fig. 2). A comparison of estimated abundance with estimates from the triennial survey (Fig. 3) illustrates that both models show a small decrease in abundance from 1986 to 1992 and then an increase from 1995 to 2001. However, the magnitude of this increase is significantly smaller for the SDFA model. Comparison with estimates of spawning output from the most recent stock assessment corroborate a decrease in abundance prior to 1993 and an increase after 1993, where the magnitude of increase is smaller than either the fishery or survey indices. We therefore conclude that the fishery-dependent index captures many of the same patterns as survey and assessment estimates. Remaining differences between the assessment estimates and the survey-based index of abundance may arise because our inclusion of calendar date has not entirely controlled for changes in survey timing between 1992 and 1995.

Finally, our cross-validation experiment shows that the model with both spatial variation and targeting has substantially better predictive performance than a model without spatial variation or targeting (Fig. 4; mean improvement = 0.149 log-score per datum). The model that incorporates targeting explains nearly half of this difference (i.e., mean improvement = 0.074 log-score), while the model that incorporates spatial variation explains nearly three-quarters of this difference (mean improvement = 0.031 log-score). Therefore, including spatial variation leads to a greater improvement in predictive performance than including targeting. Further exploration of spatiotemporal variation in density among species (i.e., $\mathrm{Corr}(\mathbf{d}(s, t))$) shows that petrale density is essentially uncorrelated with dover sole and is negatively correlated with the density of both thornyheads and sablefish (Fig. 5). However, covariation

Fig. 1. Simulation experiment for estimating true trends in population abundance (black line, median; shading, the interval containing 90% of simulation replicates, where the biomass trend in each replicate is rescaled to have a maximum value of one) for four species (rows) using four configurations of the estimation model (columns) representing no spatial variation or targeting (1st column), only targeting (2nd column), only spatial variation (3rd column), or both targeting and spatial variation (4th column). For each estimation model, we rescale the estimated index to have a mean of one and then depict the error (defined as the natural log of the rescaled estimate divided by the rescaled true value) for each year, where a well-performing model will have a line and shaded region near zero (shown as a dotted grey line). [Colour online.]



due to fisher targeting (i.e., Corr(**f**(*s*, *t*))) causes a negative correlation between petrale and all other species (Fig. 6). This supports our hypothesis that fishers target either petrale or other species in the offshore winter bottom-trawl fishery.

Discussion

In this study, we have decomposed the catch equation, $\mathbb{E}(C) = QFD$, for each fishing operation (e.g., trawl set or longline soak) into components representing spatiotemporal variation in fish density (*D*), small-scale tactics that allocate fishing effort (*F*), and measurement variables affecting catchability (*Q*), as well as residual variation. This decomposition implies that a spatiotemporal model of fish density can account for spatial variation in density (and therefore control for changes in the spatial allocation of fishing effort) while filtering out covariation in model residuals (as caused by fine-scale fisher tactics). We use SDFA to implement this model and provide a simulation experiment and case-study involving petrale sole to show that SDFA can provide reasonable estimates of abundance trends using multispecies fishery data, despite shifts in fishing behavior caused by variation in prices. In both a simulation experiment and a case study involving the winter petrale sole fishery off the US West Coast, accounting for coarse-scale spatial adjustments was more important (leading to a greater decrease in predictive error) than accounting for fine-scale fishing tactics. Importantly, our decomposition assumes

Fig. 2. Map of spatiotemporal variation in density for petrale sole, estimated using SDFA applied to fishery catch rate data for the winter offshore bottom-trawl fishery.

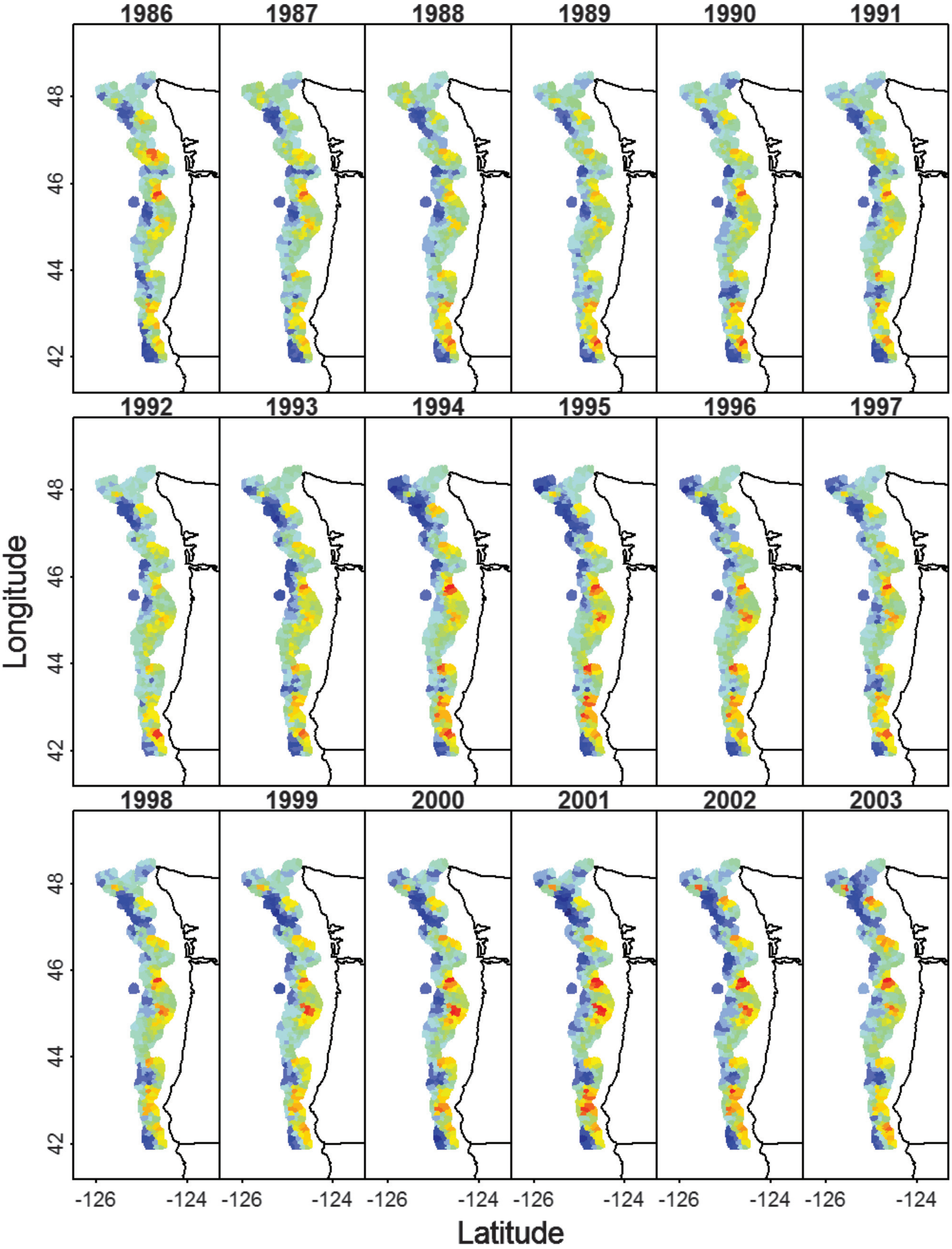
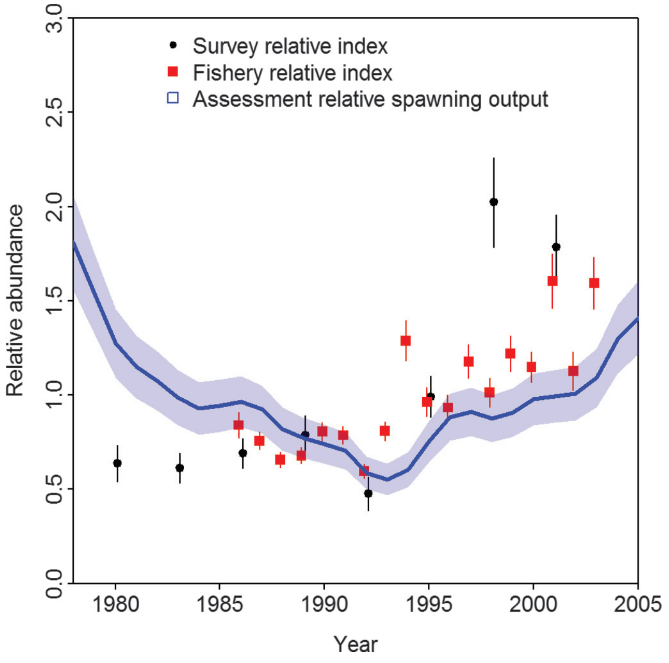


Fig. 3. Comparison of survey-based index of abundance, SDFA estimate of fishery-dependent index of abundance, and assessment estimate of spawning output for petrale sole (where each is rescaled to have a mean of one). [Colour online.]



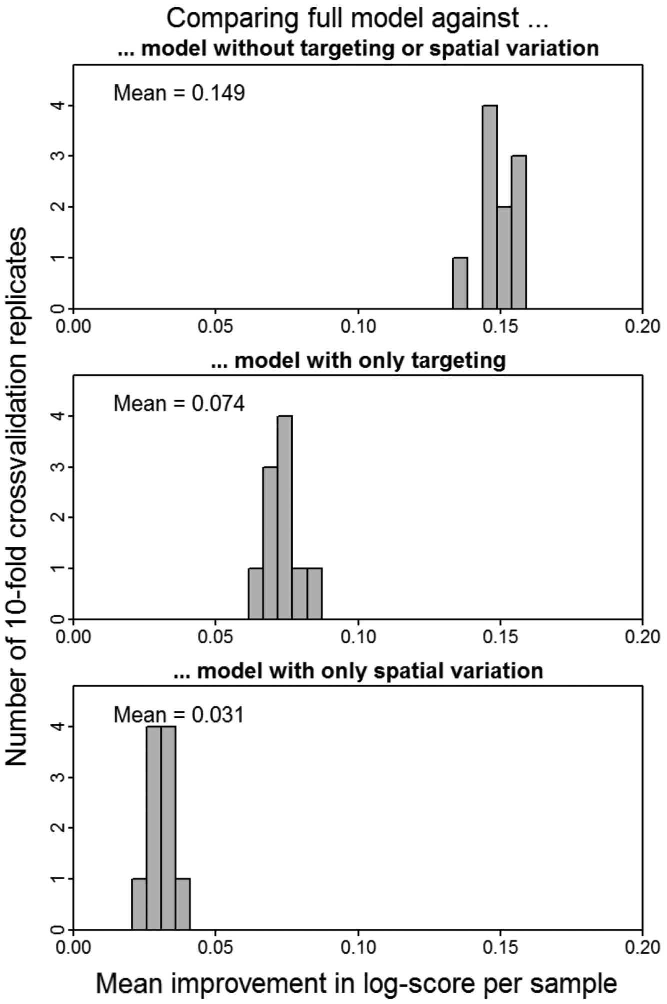
that latent variables representing fine-scale fisher tactics are orthogonal to latent variables representing density (i.e., that fishers make independent decisions about where and how to fish). We therefore hypothesize that SDFA will perform well when either (*i*) the spatial allocation of fishing effort contributes greater variance in catch rates than fine-scale tactics or (*ii*) large-scale spatiotemporal variation and fine-scale tactics are statistically independent. However, the relative variance of multispecies catch rates caused by large-scale spatial decisions vs. small-scale tactics remains an open research question (Abbott et al. 2015), and relatively little is known about mechanisms for species targeting in many fisheries. Therefore, we cannot confidently speculate about when these two conditions will be met in real-world fisheries.

Spatiotemporal methods are increasingly used to estimate fish density in fisheries models (Thorson et al. 2015*a*). Research suggests that spatiotemporal models are more statistically efficient (e.g., have a lower expected root-mean-squared error) than spatially stratified models (see definitions in section titled "Study goals") in some cases (Shelton et al. 2014; Thorson et al. 2015*c*). When using spatiotemporal models to predict density, the estimated variance of density predictions is generally greater at locations that are geographically distant from available data than locations that are close to available data (e.g., as seen for northern dusky rockfish in the Gulf of Alaska; Lunsford et al. (2015); D. Hanselman, personal communication, 2015). We believe that this is an appropriate way to characterize uncertainty when sampling is not distributed evenly throughout space, so we consider this characteristic to be a benefit of index standardization using spatiotemporal models. Spatially stratified models remain more common than spatiotemporal models when analyzing fishery CPUE data (Carruthers et al. 2011; Ono et al. 2015; Campbell 2016).

SDFA also leverages multispecies information in a way that is relatively new in fisheries and ecology. Joint species distribution models are increasingly proposed as a "Robin Hood" approach (Punt et al. 2011) for estimating species distributions, i.e., a way to borrow information from well-surveyed species to inform estimates of density for infrequently encountered species (Warton et al. 2015; Thorson et al. 2015*b*; Ovaskainen et al. 2016). However, we know of no previous fisheries study that has developed a statistical model to jointly estimate changes over time in the distribution of multiple species (termed a "joint dynamic species distribution model", JDSDM). In addition to estimating changes in total abundance, JDSMDs can be used to track changes in population range over time, as well as to estimate the total area of the core population range (Thorson et al. (2016*a*)). Tracking both total (or relative) abundance and core area could potentially improve monitoring for cases in which stocks radically shift their density in their core range, as has been implicated in both the collapse and recovery of northern cod (Walters and Maguire 1996). The ability to estimate shifts in species range using fishery-dependent catch rates would also greatly expand the data available to monitor climate impacts on marine fish distribution.

Despite these benefits, estimates of abundance using a spatiotemporal approach will likely be biased if (*i*) large areas of the population domain have no fishery data and are therefore unobserved and (*ii*) trends in the unobserved areas are systematically different from those in the observed areas (see, e.g., Campbell (2016) for a discus-

Fig. 4. Improvement in log-predictive probability per observation from a 10-fold cross-validation experiment for a "full" model (which includes multispecies information, i.e., estimating **A** and **B**, where $J = K = P$) relative to three alternative models: (top panel) neither targeting nor spatial variation (**A** = **B** = **0**); (middle panel) only small-scale targeting (**A** = **0**, estimating **B**); or (bottom panel) only spatial variation (**B** = **0**, estimating **A**). Each panel also lists the mean improvement in log-probability per observation across all 10 cross-validation replicates in the upper-left of each panel.



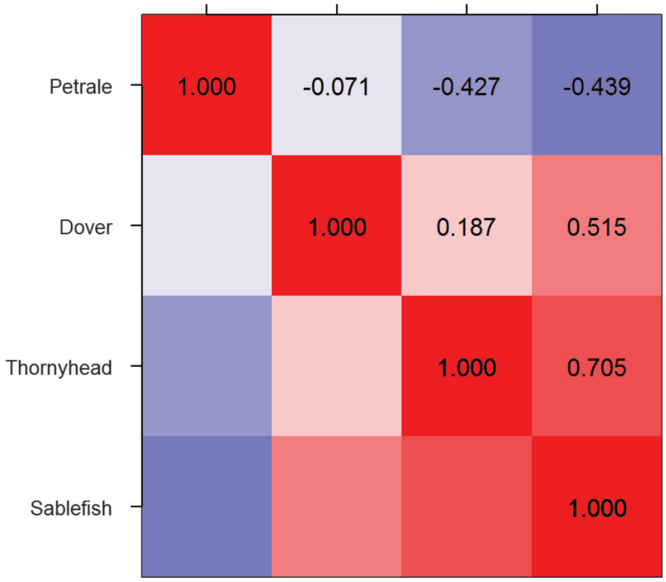


Fig. 5. Spatiotemporal correlation in density among species, Corr($\mathbf{d}(s, t)$). [Colour online.]

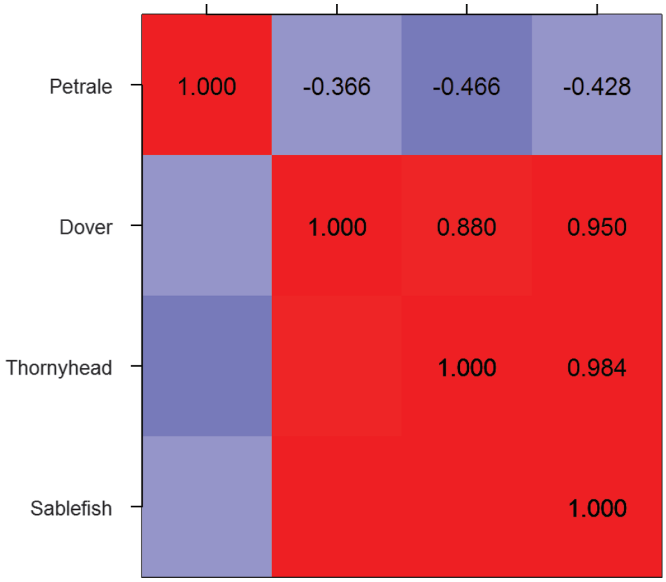


Fig. 6. Spatiotemporal correlation in fishing effort caused by small-scale targeting among species, Corr($\mathbf{f}(s, t)$). [Colour online.]

sion). In this case, the SDFA model developed here will imply the assumption that trends are similar between observed and unobserved areas (with greater predictive variance in unobserved areas), and this is a poor assumption in some cases, e.g., during spatial expansion of a fishing fleet (Walters 2003; Kleiber and Maunder 2008). However, we note that the SDFA model could potentially be modified to account for spatial expansion of a fishery. For example, if some spatial areas are entirely unobserved during early years (e.g., during the development of a fishery), then SDFA can be configured such that local log-densities follow a random-walk process (e.g., that $\rho = 1$ and $\boldsymbol{\omega}_k = 0$). In this case, SDFA estimates of density prior to the first year of fishing will be equal to density in the first year. This specification is similar to the "imputation" method of Carruthers et al. (2010), and simulation testing has shown that it decreases bias when imputing density in unobserved areas relative to simple design-based approaches (Thorson et al. (2016*b*)).

As with any new statistical method, we recommend several avenues of future research to follow up on the research presented here.

1. We envision that the biggest improvements will arise from incorporating additional covariates to explain spatial variation in either density or catchability. Including habitat variables has been shown to improve density predictions in spatial models (Shelton et al. 2014) and could also improve predictions in unobserved areas whenever there is a strong link between a habitat variable and population density (e.g., Lehodey et al. 1997). Similarly, including variables affecting catchability might be useful both to distinguish trends in fishing power over time (Robins et al. 1998) and to predict small-scale tactics for individual fishing operations (Bigelow and Maunder 2007). In particular, we note that individual quota-based management systems provide a natural experiment for estimating the potential impact of fine-scale tactics. In this case, the quantity of quota remaining for target and nontarget species could be included as a catchability covariate, and this would provide contrast to estimate the degree to which fishers can influence the species composition of their catch to avoid choke species (Branch et al. 2006; Abbott et al. 2015).
2. We also acknowledge that many regions do not collect detailed spatial information regarding fishery catches and that available spatial records might not be accurate. For example, the shark fishery off Southeast Australia reports fishery catches within relatively coarse strata, and this has limited the potential for spatial analysis of these data (Punt et al. 2000). We therefore recommend that future studies explore the impact of missing, coarse-resolution, or erroneous spatial information on the performance of SDFA. In the limit of no spatial information, SDFA will attribute all covariation in catches to fisher tactics (i.e., **B**). In this case, it will be unable to account for changes in the spatial distribution of fishing effort (as also seen in "Targeting" scenario in our simulation study, or Appendix C). This scenario bears some resemblance to the "direct principle components" (DPC) method of Winker et al. (2013, 2014), and we recommend further simulation comparison of nonspatial models for standardizing fishery catch rate data under different data-generating models (in particular the DPC method). We also envision continued research regarding best practices for when to estimate a nonlinear relationship between an estimated index of abundance and stock biomass (Harley et al. 2001; Wilberg et al. 2010), which will likely be necessary when using a nonspatial model for index standardization.
3. By default, SDFA estimates that loadings parameters (**A** and **B**) are constant over time. However, future studies could explore either random-walk or time-blocked approaches to changes in spatiotemporal covariation or fisher tactics over time. For example, small-scale fisher targeting likely underwent a gradual change in the petrale fishery following the implementation of individual quotas in 2003–2004. If future studies jointly analyze data from both before and after the introduction of individual quotas, we envision that the covariance of fishery tactics (**B**) would be estimated with a time block in that year. Similarly, changes in ecosystem dynamics (i.e., the introduction or recovery of a new predator) might lead to changes in spatiotemporal covariation in density (Swain et al. 2015).

Despite these several topics for future research, we believe that SDFA offers a valuable addition to the toolbox for analyzing fishery-dependent catch rates. In particular, we suggest that SDFA is more easily interpreted than previous two-stage approaches to incorporating fisher tactics during index standardization (e.g., Stephens and MacCall 2004). Interpretability provides several benefits, including (*i*) allowing insight about which variables are estimable or confounded when included as density or catchability covariates, (*ii*) permitting the testing of hypotheses and interpretation of estimates regarding the covariance between species in density or fisher tactics, and (*iii*) allowing inspection of model residuals and application of model selection methods.

Acknowledgements

We thank R. Hovel and M. Scheuerell for insights leading to the development of spatial dynamic factor analysis, and as always, we thank K. Kristensen for developing Template Model Builder, without which this project would not be possible. We also thank M. Bravington and A. Haynie for discussions regarding fisher targeting and A. Berger, D. Goethel, and P. Lynch for organizing the symposium entitled "Space Oddity: Recent Advances Incorporating Spatial Processes in the Fishery Stock Assessment and Management Interface", occurring at the 2015 National AFS Meeting in Portland, Oregon. K. Ono was partially funded by the Joint Institute for the Study of the Atmosphere and Ocean (JISAO) under NOAA Cooperative Agreement NA10OAR4320148, Contribution No. 2570. Finally, we thank A. Haynie and two anonymous reviewers for comments on an earlier draft.

References

Abbott, J.K., Haynie, A.C., and Reimer, M.N. 2015. Hidden flexibility: institutions, incentives, and the margins of selectivity in fishing. Land Econ. **91**(1): 169–195. doi:10.3368/le.91.1.169.

Beverton, R.J.H., and Holt, S.J. 1957. On the dynamics of exploited fish populations. Chapman & Hall, London.

Bigelow, K.A., and Maunder, M.N. 2007. Does habitat or depth influence catch rates of pelagic species? Can. J. Fish. Aquat. Sci. **64**(11): 1581–1594. doi:10.1139/f07-115.

Branch, T.A., Hilborn, R., Haynie, A.C., Fay, G., Flynn, L., Griffiths, J., Marshall, K.N., Randall, J.K., Scheuerell, J.M., Ward, E.J., and Young, M. 2006. Fleet dynamics and fishermen behavior: lessons for fisheries managers. Can. J. Fish. Aquat. Sci. **63**(7): 1647–1668. doi:10.1139/f06-072.

Campbell, R.A. 2016. A new spatial framework incorporating uncertain stock and fleet dynamics for estimating fish abundance. Fish Fish. **17**(1): 56–77. doi:10.1111/faf.12091.

Carruthers, T.R., McAllister, M.K., and Ahrens, R.N.M. 2010. Simulating spatial dynamics to evaluate methods of deriving abundance indices for tropical tunas. Can. J. Fish. Aquat. Sci. **67**(9): 1409–1427. doi:10.1139/F10-056.

Carruthers, T.R., Ahrens, R.N.M., McAllister, M.K., and Walters, C.J. 2011. Integrating imputation and standardization of catch rate data in the calculation of relative abundance indices. Fish. Res. **109**(1): 157–167. doi:10.1016/j.fishres.2011.01.033.

Deporte, N., Ulrich, C., Mahévas, S., Demanèche, S., and Bastardie, F. 2012. Regional métier definition: a comparative investigation of statistical methods using a workflow applied to international otter trawl fisheries in the North Sea. ICES J. Mar. Sci. J. Conserv. **69**(2): 331–342. doi:10.1093/icesjms/fsr197.

Gelman, A., Hwang, J., and Vehtari, A. 2013. Understanding predictive information criteria for Bayesian models. Stat. Comput. **24**(6): 997–1016. doi:10.1007/s11222-013-9416-2.

Haltuch, M.A., Ono, K., and Valero, J. 2013. Status of the US petrale sole resource in 2012. Northwest Fisheries Science Center, National Marine Fisheries Service, Seattle, Washington.

Harley, S.J., Myers, R.A., and Dunn, A. 2001. Is catch-per-unit-effort proportional to abundance? Can. J. Fish. Aquat. Sci. **58**(9): 1760–1772. doi:10.1139/f01-112.

Haynie, A.C., and Layton, D.F. 2010. An expected profit model for monetizing fishing location choices. J. Environ. Econ. Manage. **59**(2): 165–176. doi:10.1016/j.jeem.2009.11.001.

He, X., Bigelow, K.A., and Boggs, C.H. 1997. Cluster analysis of longline sets and fishing strategies within the Hawaii-based fishery. Fish. Res. **31**(1–2): 147–158. doi:10.1016/S0165-7836(96)00564-4.

Kass, R.E., and Steffey, D. 1989. Approximate Bayesian inference in conditionally independent hierarchical models (parametric empirical Bayes models). J. Am. Stat. Assoc. **84**(407): 717–726. doi:10.1080/01621459.1989.10478825.

Kissling, W.D., Dormann, C.F., Groeneveld, J., Hickler, T., Kühn, I., McInerny, G.J., Montoya, J.M., Römermann, C., Schiffers, K., Schurr, F.M., et al. 2012. Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents. J. Biogeogr. **39**(12): 2163–2178. doi:10.1111/j.1365-2699.2011.02663.x.

Kleiber, P., and Maunder, M.N. 2008. Inherent bias in using aggregate CPUE to characterize abundance of fish species assemblages. Fish. Res. **93**(1–2): 140–145. doi:10.1016/j.fishres.2008.03.013.

Kotwicki, S., De Robertis, A., Ianelli, J.N., Punt, A.E., and Horne, J.K. 2013. Combining bottom trawl and acoustic data to model acoustic dead zone correction and bottom trawl efficiency parameters for semipelagic species. Can. J. Fish. Aquat. Sci. **70**(2): 208–219. doi:10.1139/cjfas-2012-0321.

Kristensen, K. 2014. TMB: general random effect model builder tool inspired by ADMB. Available from https://github.com/kaskr/adcomp.

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.

Lehodey, P., Bertignac, M., Hampton, J., Lewis, A., and Picaut, J. 1997. El Niño Southern Oscillation and tuna in the western Pacific. Nature, **389**(6652): 715–718. doi:10.1038/39575.

Lindgren, F., Rue, H., and Lindström, J. 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. J. R. Stat. Soc. Ser. B Stat. Methodol. **73**(4): 423–498. doi:10.1111/j.1467-9868.2011.00777.x.

Lunsford, C.R., Hulson, P.-J.F., Shotwell, S.K., and Hanselman, D.H. 2015. 12. Assessment of the Dusky Rockfish stock in the Gulf of Alaska. 2015 Plan Team, Alaska Fisheries Science Center, National Marine Fisheries Service, Seattle, Washington. Available from http://www.afsc.noaa.gov/refm/stocks/plan_team/GOAdusky.pdf.

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.

Maynou, F., Demestre, M., and Sánchez, P. 2003. Analysis of catch per unit effort by multivariate analysis and generalised linear models for deep-water crustacean fisheries off Barcelona (NW Mediterranean). Fish. Res. **65**(1): 257–269. doi:10.1016/j.fishres.2003.09.018.

Methot, R.D., and Wetzel, C.R. 2013. Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fish. Res. **142**: 86–99. doi:10.1016/j.fishres.2012.10.012.

Ono, K., Punt, A.E., and Hilborn, R. 2015. Think outside the grids: an objective approach to define spatial strata for catch and effort analysis. Fish. Res. **170**: 89–101. doi:10.1016/j.fishres.2015.05.021.

Ovaskainen, O., Roy, D.B., Fox, R., and Anderson, B.J. 2016. Uncovering hidden spatial structure in species communities with spatially explicit joint species distribution models. Methods Ecol. Evol. **7**(4): 428–436. doi:10.1111/2041-210X.12502.

Pascoe, S., Hutton, T., Thébaud, O., Deng, R., Klaer, N., and Vieira, S. 2015. Setting economic target reference points for multiple species in mixed fisheries. CSIRO Oceans and Atmosphere Flagship, Brisbane, Australia.

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**(1): 51–65. doi:10.1139/f99-176.

Punt, A.E., Walker, T.I., Taylor, B.L., and Pribac, F. 2000. Standardization of catch and effort data in a spatially-structured shark fishery. Fish. Res. **45**(2): 129–145. doi:10.1016/S0165-7836(99)00106-X.

Punt, A.E., Smith, D.C., and Smith, A.D. 2011. Among-stock comparisons for improving stock assessments of data-poor stocks: the "Robin Hood" approach. ICES J. Mar. Sci. J. Conserv. **68**(5): 972–981. doi:10.1093/icesjms/fsr039.

R Core Team. 2013. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from http://www.R-project.org/.

Robins, C.M., Wang, Y.G., and Die, D. 1998. The impact of global positioning systems and plotters on fishing power in the northern prawn fishery, Australia. Can. J. Fish. Aquat. Sci. **55**(7): 1645–1651. doi:10.1139/f98-037.

Sanchirico, J.N., Holland, D., Quigley, K., and Fina, M. 2006. Catch–quota balancing in multispecies individual fishing quotas. Mar. Pol. **30**(6): 767–785. doi:10.1016/j.marpol.2006.02.002.

Schaefer, M.B. 1954. Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. Bull. Inter-Am. Trop. Tuna Comm. **1**: 27–56.

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.

Smith, T.D. 2007. Scaling fisheries: the science of measuring the effects of fishing, 1855–1955. 1st ed. Cambridge University Press, Cambridge, U.K.

Soberón, J. 2007. Grinnellian and Eltonian niches and geographic distributions of species. Ecol. Lett. **10**(12): 1115–1123. doi:10.1111/j.1461-0248.2007.01107.x.

Somers, K.A., Bellman, M., Jannot, J., Riley, N., and McVeigh, J. 2014. Estimated discard and catch of groundfish species in the 2013 US West Coast fisheries. West Coast Groundfish Observer Program, National Marine Fisheries Service, NWFSC, Seattle, Washington.

Stephens, A., and MacCall, A. 2004. A multispecies approach to subsetting logbook data for purposes of estimating CPUE. Fish. Res. **70**(2): 299–310. doi:10.1016/j.fishres.2004.08.009.

Swain, D.P., Benoît, H.P., and Hammill, M.O. 2015. Spatial distribution of fishes in a Northwest Atlantic ecosystem in relation to risk of predation by a marine mammal. J. Anim. Ecol. **84**(5): 1286–1298. doi:10.1111/1365-2656.12391.

Thorson, J.T., and Minto, C. 2015. Mixed effects: a unifying framework for statistical modelling in fisheries biology. ICES J. Mar. Sci. **72**(5): 1245–1256. doi:10.1093/icesjms/fsu213.

Thorson, J.T., and Ward, E. 2013. Accounting for space–time interactions in index standardization models. Fish. Res. **147**: 426–433. doi:10.1016/j.fishres.2013.03.012.

Thorson, J.T., Skaug, H.J., Kristensen, K., Shelton, A.O., Ward, E.J., Harms, J.H., and Benante, J.A. 2014. The importance of spatial models for estimating the strength of density dependence. Ecology, **96**(5): 1202–1212. doi:10.1890/14-0739.1.

Thorson, J.T., Ianelli, J.N., Munch, S.B., Ono, K., and Spencer, P.D. 2015*a*. Spatial delay-difference models for estimating spatiotemporal variation in juvenile production and population abundance. Can. J. Fish. Aquat. Sci. **72**(12): 1897–1915. doi:10.1139/cjfas-2014-0543.

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

Thorson, J.T., Ianelli, J.N., Larsen, E., Ries, L., Scheuerell, M.D., Szuwalski, C., and Zipkin, E. 2016*a*. Joint dynamic species distribution models: a tool for community ordination and spatiotemporal monitoring. Glob. Ecol. Biogeogr. [In press.] doi:10.1111/geb.12464.

Thorson, J.T., Pinsky, M.L., and Ward, E.J. 2016*b*. Model-based inference for estimating distribution changes in marine species. Methods Ecol. Evol. [In press.] doi:10.1111/2041-210X.12567.

Walters, C. 2003. Folly and fantasy in the analysis of spatial catch rate data. Can. J. Fish. Aquat. Sci. **60**(12): 1433–1436. doi:10.1139/f03-152.

Walters, C., and Maguire, J.-J. 1996. Lessons for stock assessment from the northern cod collapse. Rev. Fish Biol. Fish. **6**(2): 125–137. doi:10.1007/BF00182340.

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. **30**(12): 766–779. doi:10.1016/j.tree.2015.09.007. PMID:26519235.

Weinberg, K.L. 2002. The 2001 Pacific West Coast bottom trawl survey of groundfish resources: estimates of distribution, abundance, and length and age composition. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Alaska Fisheries Science Center.

Wilberg, M.J., Thorson, J.T., Linton, B.C., and Berkson, J. 2010. Incorporating time-varying catchability into population dynamic stock assessment models. Rev. Fish. Sci. **18**(1): 7–24. doi:10.1080/10641260903294647.

Winker, H., Kerwath, S.E., and Attwood, C.G. 2013. Comparison of two approaches to standardize catch-per-unit-effort for targeting behaviour in a multispecies hand-line fishery. Fish. Res. **139**: 118–131. doi:10.1016/j.fishres.2012.10.014.

Winker, H., Kerwath, S.E., and Attwood, C.G. 2014. Proof of concept for a novel procedure to standardize multispecies catch and effort data. Fish. Res. **155**: 149–159. doi:10.1016/j.fishres.2014.02.016.

Appendix A: Description of the simulation model

We use a simulation experiment to explore the performance of the proposed spatial dynamic factor analysis model (SDFA) when estimating an index of abundance using multispecies fishery catch rate data. The experiment involves simulating realistic data with known properties using a simulation model and then fitting SDFA to compare SDFA estimates with known properties from the simulation model. To generate data, we use a simulation model that tracks population abundance for *P* species in each of *S* grid cells over *T* time intervals. Code to use this simulation model is available using function Generate_scenario_data in the Spatial_DFA R package on the first author's GitHub page (https://github.com/James-Thorson/spatial_DFA), and it is a generalization of the simulator used by Ono et al. (2015) that was developed independently of the SDFA model presented in the main text.

Population dynamics

In this simulation model, the abundance of species p in grid cell s and time t follows a logistic population growth function with movement dynamics controlled by the species-specific depth preference and movement rates. The population density, $d_p(s, t)$, for species p in grid s at time t changes through time as a function of the catch, $c_p(s, t)$, the maximum per-capita population growth rate, r_p, the carrying capacity of species p in cell s, $k_p(s)$, and the proportion of individuals emigrating and immigrating:

$$d_p^*(s,t+1) = d_p(s,t) + r_p d_p(s,t)\left(1 - \frac{d_p(s,t)}{k_p(s)}\right) - c_p(s,t) - \sum_{s^*=1}^{S} m_p(s^*,s) d_p(s,t) + \sum_{s^*=1}^{S} m_p(s,s^*) d_p(s^*,t)$$

The probability $m_p(s^*, s)$ of species p moving from grid s to grid s^* is a function of

1. the mobility of the species s, where the probability of moving between cells s and s^* is modeled as an negative-exponential function of distance between cells and a species-specific parameter λ_p representing mean movement rates;
2. species depth preferences, where each species has a preferred depth, $u_p(z)$, and habitat preferences decline away from this preferred depth following a lognormal distribution, where $\sigma_p(z)$ governs the rate at which habitat preferences decline away from preferred depths. Depth $z(s)$ for each cell s is simulated as a Gaussian random field with an pointwise variance σ_z^2, where the covariance between cells follows a negative exponential decay that depends on distance between grid cells and where $z(s)$ is added to a constant to ensure that it remains positive; and
3. species latitudinal preferences, where each species has a preferred latitude, $u_p(x)$, and habitat preferences decline away from this preferred latitude following a lognormal distribution, where $\sigma_p(x)$ governs the rate at which habitat preferences decline away from preferred latitude.

Movement rates from cell s^* to cell s is then affected by these three behaviors:

$$m_p(s^*,s) = c_s^{-1} \times \exp(-\lambda_p \Delta s) \times z(s^*)^{-1} \exp\left(\frac{\left(\log(z(s^*)) - \log(u_p(z))\right)^2}{2\sigma_p(z)}\right) \times \exp\left(\frac{\left(x(s^*) - u_p(x)\right)^2}{2\sigma_p(x)}\right)$$

where c_s^{-1} is an integration constant defined so that each column of the movement matrix $\mathbf{M}_p$ has columns that sum to one.

Species biomass at the start of the simulation is assumed to be at carrying capacity $k_p(s)$ for each cell s. Carrying capacity is calculated by determining the stationary distribution of population density for all cells given productivity and movement parameters.

Effort dynamics

The total number of fishing trips, $U(t)$, for all vessels during year t is generated from a discretized lognormal distribution with mean given by a logistic function of time and a coefficient of variation of 0.2:

$$\begin{aligned} U^*(t) &\sim \text{Normal}\left(\frac{U_T}{1 + \exp(-0.1t)}, 0.2\right) \\ U(t) &= \lfloor \exp(U^*(t)) \rfloor \end{aligned}$$

where U_T is a parameter representing the level of fishing effort in the final year (where we use $U_T = 100$, and $\lfloor \exp(U^*(t)) \rfloor$ takes the value of $\exp(U^*(t))$ and rounds it down. The vessel conducting trip u of the $U(t)$ trips in year t is drawn from a uniform categorical distribution:

$$v(u) \sim \text{Categorical}(\boldsymbol{\pi}^*)$$

where Categorical($\mathbf{x}$) is a categorical distribution in which the probability of category i is x_i, and $\boldsymbol{\pi}^*$ is a vector with length V, where each element $\pi_v^* = V^{-1}$. The probability of fishing in cell s in year t depends on the mean expected revenue in each cell given density $d_p(s, t)$ and price $\phi_p(t)$ of species p in year t:

$$\pi(s,t) = \frac{\sum_{p=1}^{P} \phi_p(t) d_p(s,t)}{\sum_{s=1}^{S} \sum_{p=1}^{P} \phi_p(t) d_p(s,t)}$$

Trip u is then assigned to cell s following a categorical distribution:

$$s(u) \sim \text{Categorical}(\boldsymbol{\pi}(t))$$

where $\boldsymbol{\pi}(t)$ is a vector with S elements $\pi(s, t)$.

Given this allocation of fishing trips among vessels and grid cells, we then simulate fishery catch for each species and trip. Each vessel has variation in mean catchability, q_v, generated from a lognormal distribution:

$$\log(q_v) \sim \text{Normal}(\log(0.2), 0.2)$$

and catchability $q_p(u)$ for species p on trip u follows a Tweedie distribution:

$$q_p(u) \sim \text{Tweedie}(q_{v(u)}, 1.2, 0.2)$$

where Tweedie(x, y, z) is a Tweedie distribution with expectation x, power y, dispersion z, and a variance of $0.1q_v^{1.2}$. Total catch $c_{p,t}(u)$ for trip u of species p in time t is then calculated as

$$c_{p,t}(u) = d_p(s(u), t)(1 - \exp(-q_p(u)))$$

Total catch $c_p(s, t)$ for each species p, cell s, and year t is used in the population dynamics component and is calculated as

$$c_p(s, t) = \sum_{u=1}^{U(t)} I(s(u) = s)c_{p,t}(u)$$

where $I(s(u) = s)$ is an indicator function that is one if trip u occurs in cell s (i.e., $s(u) = s$) and zero otherwise. The spatial dynamic factor analysis (SDFA) model then receives a data frame with a row for each fishing trip and columns representing the year (t), cell $s(u)$, and catch $c_{p,t}(u)$ for each species p.

Reference

Ono, K., Punt, A.E., and Hilborn, R. 2015. Think outside the grids: an objective approach to define spatial strata for catch and effort analysis. Fish. Res. **170**: 89–101. doi:10.1016/j.fishres.2015.05.021.

Appendix B

Table B1. Parameter values for each species used in the simulation experiment.

Name	Symbol	Species 1	Species 2	Species 3	Species 4
Initial price of species p in year t	$\phi_p(1 \le t < 10)$	1	1	1	1
Middle price of species p in year t	$\phi_p(11 \le t < 20)$	2	1	0.5	0.5
Final price of species p in year t	$\phi_p(21 \le t < 30)$	1	1	1	1
Movement rate	λ_p	0.4	0.4	0.4	1.0
Preferred depth	$\mu_p(z)$	220	450	250	1200
Depth range	$\sigma_p(z)$	0.5	0.5	0.3	0.5
Preferred latitude	$\mu_p(x)$	0	0	0	0
Latitudinal range	$\sigma_p(x)$	50	50	50	50
Total initial abundance	$\sum_{s=1}^{S} d_p(s,1)$	50 000	820 000	36 000	540 000
Maximum per-capita population growth	r_p	0.22	0.17	0.08	0.08

Appendix C: Sensitivity of simulation experiment to spatial resolution

We also present results when repeating the simulation experiment but using a coarse resolution for spatial data. Specifically, we simulate data at each of 100 locations on a 10 cell × 10 cell grid. However, we then use a *k*-means algorithm to identify 20 sites and aggregate each grid cell to the nearest site. In this way, we decrease the spatial resolution of simulated data, so that each site includes data from 5 “sites” on average.

This exercise illustrates that decreased spatial resolution for available data has no impact on models that neglect spatial variation (Fig. C1, 1st and 2nd columns). However, the performance of the spatial or the spatial and targeting models (Fig. C1, 3rd and 4th columns) are both decreased. In particular, these spatial models have performance that is more similar to the nonspatial models, with a distinct shift in bias at years 10–11 and 20–21. Biases change at these years due to changes in fish prices and resulting fisher targeting decisions.

Fig. C1. Simulation experiment for estimating true trends in population abundance (see Fig. 1 caption for details), given a coarse spatial resolution for spatial variation. [Colour online.]