Evaluating drivers and predictability of catch composition in a highly mixed trawl fishery using stacked and joint species models

James A. Smith[1*], Daniel D. Johnson[1]

[1] Port Stephens Fisheries Institute, New South Wales Department of Primary Industries,
Locked Bag 1, Nelson Bay, NSW 2315, Australia
[*] james.a.smith@dpi.nsw.gov.au

Abstract

Evaluating drivers and the predictability of catch is valuable for the management of mixed fisheries. Drivers can represent or help to identify levers for management and predictable catch compositions are a key component of simulation tools and dynamic management strategies. But modelling mixed fisheries can be challenging due to the large number of taxa, and analysis typically focuses on a few key species or highly aggregated taxa.

Here we employ seven types of stacked and joint species models to explore the drivers and predictability of trawl-level catches in an ocean prawn trawl fishery, in New South Wales, Australia. Catch data was sourced from an observer program, with 130 taxa able to be modelled. The main drivers of catch composition were latitude, depth, and seasonality represented here by water temperature. Water column mixing, lunar illumination, and fishing effort were also important for some taxa. Up to 60-80 taxa had good predictive skill (AUC>0.8, >35% decline in mean absolute error relative to an intercept-only model), and an additional 40-60 taxa had lower but still useful predictive skill (AUC>0.7, 25-35% decline in error). However, the number of skilful taxa varied considerably among model type.

The best framework for prediction was stacked random forests using a hurdle modelling approach, followed by a spatial joint species model. Our results show that predictive models at a fine spatial-temporal and taxonomic resolutions can be a viable information tool for highly mixed fisheries, but these tools ultimately need to be tested against specific management objectives and performance metrics, such as spatial closures and bycatch:target catch ratios.

Keywords: species distribution modelling, SDM, bycatch, penaeid, spatial management

1. Introduction

Highly mixed trawl fisheries are often subject to strict management to monitor and manage the high number of species caught, included discarded bycatch (Kennelly 1995, Johnsen and Eliasen 2011, Catchpole et al 2017, Kennelly 2020). Discard of an organism results from two processes: 1) being caught, and 2) being discarded once caught, and the drivers of the two processes are likely to be quite different. Here we focus on the first component – catch composition – by identifying the drivers of catch, and the level of accuracy with which we can predict trawl catches. Predicting these catches is useful for management tools such as dynamic spatial management which require finely resolved information on species catches (Dunn et al 2016, Robert et al 2019, Pons et al 2022, Panzeri et al 2023). If there is generally low predictability of the 'where and when' of catch, then this indicates that dynamic management is unlikely to be a key management tool, and instead tools such as effort control and gear modifications (Broadhurst 2000, Broadhurst et al 2012, Poos et al 2010) will be more successful, plus those which alter the second process – being discarded once caught – such as market-related levers like marketability.

Studies modelling patterns in catch often focus on key species, aggregated groups, or summary variables (e.g. bycatch, or bycatch:target catch) (Dolder et al 2018, Barnes et al 2022, Soto et al 2023), and rarely on the many taxa caught at a fine taxonomic resolution (but see for example Roberts et al 2022). Given the rapid and diverse development of multi-species modelling (Thorson et al 2016, Ovaskainen et al 2017, Zurell et al 2020), we also wanted to evaluate and compare different tools for modelling and predicting diverse catch compositions. We included three general modelling frameworks (generalized additive models, random forests, and latent variable joint generalised linear models) covering two approaches to multi-species modelling (stacked species models, and joint species models), with a goal to identify the tools best suited to predicting catch compositions for a highly mixed fishery. This fishery is the New South Wales (NSW) ocean prawn trawl (OPT) – a sector of the Ocean Trawl Fishery – which interacts with many taxa and whose discards can exceed retained catch (Kennelly et al 1998, Barnes et al 2022). The OPT makes an ideal case study for a multi-model evaluation of catch composition, due to the requirements for management and threat assessment (DPI 2007, Fletcher and Fisk 2017, DAWE 2021), and because a recent observer program has allowed a detailed accounting of all taxa caught (Johnson and Barnes 2023).

Our specific aims were to: 1) identify some key drivers of catch composition in OPT trawls at fine taxonomic and spatial-temporal scales; 2) quantify the predictive performance of catch composition in trawls, relate this to species rarity, and identify which modelling approach performs best for predicting species occurrence and biomass; and 3) identify the potential for catch composition prediction as a management tool for the OPT, especially with respect to discarded bycatch.

2. Materials and Methods

2.1. Fishery, catch data, and taxa

The Ocean Prawn Trawl fishery (OPT) of NSW is a sector of the NSW Ocean Trawl Fishery, and targets mainly Eastern king prawn (*Melicertus plebejus*), although numerous species are retained (Johnson and Barnes 2023). Typical of trawl fisheries, the diversity of species in the OPT is high (Kennelly et al 1998, Barnes et al 2022). Vessels in the OPT typically use triple-rigged trawl gear, and a fishing trip typically occurs at night within 3-30 km of shore, with a single trip comprising 1-5 individual trawls (Johnson and Barnes 2023). An average trawl fishes an area of 0.44 (0.12-0.85 95%) km^2.

We used observer data of the OPT for this analysis, because these are the only data that includes accurate locations of catches and contains records of discarded species. The observer data was collected in 2017-2019, and after minor cleaning contained 1387 trawls recorded over 421 trips, 29 vessels, and contained observations from all calendar months. Biomass of taxa caught by the OPT was estimated by trained observers by weighing each taxon. For large catches, a fixed number of each taxon was subsampled, and total biomass was estimated using subsample weight and weight of the total catch. For more information on the observer survey see Johnson and Barnes (2023).

There were 411 taxa identified in the observer program, and 136 (33%) of the taxa were recorded by observers at a coarser taxonomic resolution than species. The 411 initial taxa were too numerous to model effectively, especially given the large number of very rare taxa. To create the data for model fitting, we: 1) aggregated taxa that were rare, taxonomically related, and were discarded at similar rates in the fishery (219 taxa); and then 2) excluded rare taxa from this aggregated data set (those appearing in fewer than 30 trawls, n=89). This led to a final data set of 130 taxa, 50 at species level and 80 at a coarser level, which are listed in Table A1.

2.2. Modelling approach

Our analysis focuses on predicting trawl-level catch compositions because individual trawl is the unit of observation most relevant to understanding species interactions and catch drivers, and the most relevant to fine-scale management efforts such as spatial targeting (Dolder et al 2018) and real-time spatial management (Dunn et al 2016). We also focus on the finest feasible taxonomic resolution because this allows us to inform species-level management but in the context of the diverse trawled assemblage.

We used a suite of model types to evaluate predictive skill of catch composition in the OPT fishery because each model type has advantages and disadvantages for this task. We used generalized additive mixed models (GAMMs, Wood et al 2017), random forests (Breiman 2001), and latent variable joint generalised linear models, a type of joint species distribution model (JSDM, Warton et al 2015, Ovaskainen et al 2017). The GAMMs were included due to their flexibility and interpretability. GAMMs allow for non-linear relationships between taxon abundance and covariates, are excellent at modelling spatial patterns, and report standard regression results such as p-values and goodness-of-fit. The random forests were included for their flexibility for non-linear responses (greater than GAMMs) and potential to model rare taxa. Random forests are distribution free and can model complex patterns between taxon abundance and covariates, including interactions, but are less easy to interpret and typical study design elements cannot be specified (e.g. random effects). Random forests can also leverage techniques such as class balancing to improve disproportionate class values (here, presences and absences) which can improve the prediction of rare taxa (Stock et al 2020, Smith et al 2021). The latent variable JSDM was included because it allowed the additional modelling of species associations, which can aid the prediction of rare species (Hui et al 2013, Thorson and Barnett 2017). Additionally, JSDMs can incorporate spatial random effects, phylogenetic relatedness, and species traits into the model fitting procedure, all of which can potentially improve predictive power through ‘borrowing’ of information among taxa (Ovaskainen and Abrego 2020). However, the JSDM used here (‘HMSC’; Ovaskainen et al 2017) is based on a Bayesian GLM framework, meaning that flexibility of non-linear responses is lower than the other model types, and that model fitting is very slow which makes extensive evaluation cumbersome.

The GAMMs and random forests were used in a ‘stacked’ species distribution modelling (SDM) approach for multi-species data, and the latent variable JSDM uses a ‘joint’ SDM approach for multispecies data (Ferrier and Guisan 2006). The stacked SDM approach fits independent models to each species, then sums their predictions to estimate the community per observation unit (e.g. the catch in a single trawl). The joint SDM approach fits a model to all species at once, and can estimate residual correlations among taxa. The joint model also acknowledges that species interactions are a key part of structuring an observed assemblage. Furthermore, the JSDM used here includes latent variables to induce the species correlations, and these variables can then be used to explore the extent of unidentified ecological processes (Ovaskainen and Abrego 2020). Although a stacked SDM approach allows a different set of covariates to be included for each species, we wanted to explore performance using a single model for all species, which would be the most practical scenario when dealing with > 100 species. Stacked and joint models have previously been shown to be similarly accurate for multispecies modelling (Zurell et al 2020).

A hurdle (or ‘delta’) modelling method was included for each model type (Maunder and Punt 2004, Zuur et al 2009). This method splits the biomass data into a presence-absence component (only 1s and 0s), and an abundance-only component (only > 0 values), thus modelling separately the encounter probability and the positive catch rates. Total catch rates are then calculated by multiplying the fitted or predicted values from the two components. Hurdle methods were used for the latent variable JSDM because this was the most accurate way to model biomass within that modelling framework. Hurdle models were used for GAMMs and random forests to allow a useful comparison of methods, and because a hurdle random forest allowed us to include class balancing in the binary component to better model rare species. This class balancing used a combination of downsampling and synthetic minority oversampling (SMOTE) to ensure similar class proportions (Stock et al 2020; see *Model details* in Appendix A). We also included a single-distribution regression for the GAMMs and random forests to compare approaches. This was especially useful for the GAMM which could also use a Tweedie distribution, which provides a useful comparison to the hurdle model because, unlike the Tweedie, the hurdle assumes independent processes for its two model components. Thus, our analysis compares seven models: GAMM, GAMM hurdle, random forest, random forest hurdle, random forest hurdle with class balancing, JSDM with latent variables, JSDM with spatial latent variables, taxonomic traits, and phylogeny.

All analyses were done in R (R Core Team 2023). The GAMMs were fitted using the ‘mgcv’ package (Wood 2017), random forests with the ‘randomForest’ package (Liaw and Wiener 2002), and the JSDM was fitted using the ‘Hmsc’ package (Tikhonov et al 2020). Further details on model fitting are in Appendix A. R code is available at https://github.com/smithja16/MultiSpeciesModels_Observer.

2.3. Model covariates

Model covariates were identified based on expert opinion and previous research on trawl catch composition (Barnes et al 2022, Wang et al 2020), and the final suite of covariates were chosen using a model selection process with single-species GAMs. The single-species GAMs used different combinations of likely covariates and the most robust model was the model with the lowest AIC. The final model contained: latitude (decimal degrees), seafloor depth (fathoms), sea-surface temperature (SST, C), mixed layer depth (m), lunar illumination (proportion), and the area fished per trawl (m^2); and each variable was continuous. Calendar month was excluded due to strong collinearity with SST. A vessel covariate was included to account for correlation in trawls made by the same vessel. This was specified as a random effect in the GAMMs and JSDMs but excluded from the random forests (otherwise the model would make predictions for known vessels only). Values for these covariates were sourced from both observers and model products and are detailed in Table A2. Sources of values used when predicting to new habitats, used here for a cross validation procedure and to create example maps of species distributions, are also detailed in Table A2. The two traits included in the JSDM were asymptotic length (log(cm)) and a general habitat classification (i.e. benthic, demersal, reef, pelagic). The sources of these traits, and the creation of the taxonomic tree used in the JSDM, are detailed in Appendix A.

2.4. Evaluating drivers of catch composition

Our use of three model types as well as stacked and joint models created a valuable ensemble to conduct model-based inference to explore key drivers among taxa. The latent variable JSDM provided additional opportunity to explore residual species correlations and variable importance through the partitioning of variance. Goodness-of-fit of our models was evaluated using deviance explained (GAMMs), area under the receiver-operator curve (AUC) for classification, probit, or logit models, and otherwise R^2. Driver importance was determined using statistical significance (GAMMs), relative importance based on the mean decrease in accuracy using permutation of out-of-bag data (random forests), and posterior support values (similar to p-values; JSDMs). For the JSDMs, variance partitioning was used to help evaluate predictor importance, including the random effects. Maps of some key taxa were also predicted for an example date to highlight the differences in modelled responses to especially Latitude and Depth.

2.5. Evaluating predictive performance

We used repeated k-folds cross-validation to evaluate model predictive skill. Cross-validation evaluates model performance by fitting a model to a subset of data (a training set) and measuring how well this model predicts observations withheld from the training data (a test set). In our study we used $k = 5$ folds (i.e. 5 random splits into training and test sets), meaning that the models were trained on 80% of the data and tested against a withheld 20%. We repeated this three times to derive more accurate estimates of mean performance. For the JSDM, cross-validation was done using Hmsc's 'pcomputePredictedValues' function to greatly reduce computation time, and was not repeated but relied on having three MCMC chains to provide repetition.

Performance metrics were derived for each species by comparison of the observed and predicted test data. We calculated two main metrics: AUC and relative mean absolute error (RMAE). We use AUC and RMAE to estimate 'discrimination' and 'accuracy' respectively, i.e. good predictive performance in our study means we correctly identify trawls with higher and lower catches of a taxa (discrimination), and we can accurately estimate the biomass of a taxon in the catch (accuracy). AUC was calculated for the hurdle models only, using their binomial component. RMAE is calculated as the mean absolute difference between observed and predicted values, normalized by dividing by the mean of observed values. Mean absolute error was used instead of root mean square error, because the latter places more weight on outliers, which is less likely to make a fair comparison against an intercept-only baseline (see below). AUC was calculated using the pROC R package (Robin et al 2011) or within the Hmsc package.

For AUC, a value of 0.5 indicates a model that predicts presences and absences at random, and under 0.7 is typically considered poor predictive performance. To identify values of

RMAE that indicate good or poor performance, for each taxon we compared RMAE from their full model with the RMAE when using the mean overall catch of that taxon as predicted values (i.e. an intercept-only model). Our metric then became the decline in median RMAE as a proportion of the intercept-only RMAE (indicated by ‘RMAEp’). Using this baseline was important because species prevalence can influence the RMAE but not necessarily represent a change in performance. A model which is poor at predicting species abundance will tend to be no better than an intercept-only model (an intercept-only model will also generate an AUC around 0.5).

3. Results

3.1. Model diagnostics

Goodness of fit was generally moderate to high across model types. The GAMMs had a mean explained deviance of 34 to 51%, the random forests had a mean R^2 around 0.85 (but machine learning methods are best evaluated with out-of-sample performance), and the JSDMs had a mean AUC around 0.9 and a mean R^2 of 0.5. These goodness-of-fit values show the reasonably high level of information in the catch compositions and the ability of our covariates to explain it.

Residuals for the GAMMs showed the gamma distribution was better suited to fitting larger abundances and the Tweedie at fitting smaller abundances, but the gamma residuals were more normal suggesting a better model (admitting it had fewer data to fit). It was also clear that fitting the same GAMM for all taxa, especially the wiggliness (k parameter), left some information unexplained, and more complex models for data-rich taxa could improve relationships with explanatory variables. The JSDMs showed generally good convergence for the non-spatial model based on values for effective sample size and scale reduction factor, but this was less clear for the spatial model. For the presence-absence components around 20-30% of the taxa had a poor trace plot for at least one coefficient, and 30-40% for the abundance-only component. This is not unexpected for very rare or very common species, which show little variation among observations (Ovaskainen et al 2016). However, these poor trace plots and the reduced effective sample size show that increasing iterations and the thinning interval could be beneficial, although this needs to be weighed against our already long run times.

3.2. Drivers of catch composition

Catch composition was driven mainly by location, depth, and seasonality, i.e. Latitude, Depth and SST. Local water column mixing and illumination (i.e. MLD and Lunar) were also influential but for fewer taxa. Fishing effort (i.e. Area swept) was frequently the least important variable, showing that individual trawls are of a consistent level of effort to interact with most taxa. Covariate importance is most clearly seen in the random forest's relative importance metric (Fig. 1a) and the JSDM's partitioning of variance (Fig. 1b). A similar pattern was seen across all model types (Table A3). We do not focus here on the specific responses across species to each variable, because they are varied and numerous. Maps of predicted biomass for some taxa are shown in Fig. 2, illustrating some of the diverse relationships to Latitude, Depth, and the environment, as well as some differences among modelling approaches.

Fitting the JSDMs with Hmsc allowed us to explore the random effects (latent variables) in detail, including their relative importance (Fig. 1b). The Vessel identifier contributed much to the explained information, and we consider it likely this represents predominantly residual spatial and temporal autocorrelation, and possibly a vessel size effect or characteristic style of fishing. The Sample (observation-level) random effect typically represents species interactions and missing covariates. Based on the partitioning of variance, we can see that in the spatial JSDM the latent variation is around 40% but varying considerably among taxa (Fig. 1b). The residual correlations induced by the Sample random effect are also insightful about which taxa may be interacting and the direction of the residual information. The residual correlations (Figs. A1, A2) show that many taxa are found together more often than expected, and that the spatial model reduces these correlations (i.e. explains more of the information). The spatial scale of the leading spatial latent variable was larger and stronger for the abundance-only component at around 150 km (95% credibility interval 10 to 430 km) and only 20 km (0 to 460 km) for the presence component. This indicates potentially multiple structuring environmental variables at smaller and larger scales that remain unknown.

Traits were somewhat related to patterns in catch compositions, explaining 7% and 15% of the variation in the presence and abundance-only model components. Traits for body size and habitat were often correlated with Depth and Latitude; e.g. larger fish were more likely to be caught at deeper depths, and larger fish had greater catch biomass at more southern latitudes. Taxonomy was more important, with rho parameters 0.60 (95% credibility interval 0.41-0.75) and 0.87 (0.73-0.97) for the presence and abundance-only components respectively, which indicate a strong taxonomic signal in catch composition. This signal can be challenging to discern, but one example is that the Pleuronectiformes group of taxa are more likely to be present in trawls in shallow depths (Fig. A3).

3.3. Predictive performance

There were 130 taxa included in our analysis, and across most modelling approaches at least 80% had at least some predictive skill (AUC > 0.7, RMAEp > 0.15), and 25-45% had good or better predictive skill (AUC > 0.8, RMAEp > 0.25; Table 1). Exceptions were the GAMM hurdle and random forest with class balancing, which each performed poorly for one metric. We assume that the 89 taxa deemed too rare to model could be included in predictions of catch composition at a constant mean rate (i.e. an intercept-only model). However, if some of these species are of great interest, simpler models than we tested (using one or two covariates) could be informative, although evaluating predictive skill is less reliable with few records. The five taxa with highest predictive performance that are important taxa to the fishery were: stout whiting (*Sillago robusta*), two slipper lobster taxa (*Ibacus* spp.), broadbrow flounder (*Crossorhombus valderostratus*), and tiger flathead (*Platycephalus richardsoni*); all taxa are ranked by their performance in Table A4.

There was considerable variation in performance among modelling approaches (Table 1), but the random forest hurdle model with and without class balancing were best for discrimination (AUC), and the random forest hurdle and spatial JSDM were best for accuracy (RMAEp). Although class balancing improved discrimination this was at the cost of inflated biomass estimates and a decrease in accuracy (Table 1). The best model for each metric could predict 64% of all taxa (and 73% of key taxa) with good discrimination (AUC > 0.8), and 52% of all taxa (and 60% of key taxa) with high accuracy (RMAEp > 0.35). Models were often more accurate (RMAEp) as the prevalence of taxa decreased, which may be because absolute improvements in predictive performance can be proportionally larger when mean biomass is very low. The variation in predictive performance and among taxa is illustrated for the random forest hurdle model in Fig. A4.

As another illustration of our predictive performance, we show in Figure 3 observed and predicted catch compositions for an example trawl. The example trawl is from a withheld dataset in a 5-fold cross validation procedure. The extent that a prediction is wrong in terms of total biomass and species richness are useful metrics to help interpret our RMAEp values. Figure 3 shows these metrics for one sample, where the model with the lowest error was the random forest hurdle. Over many samples, and for the most common 50 taxa, the mean observed biomass and number of taxa per trawl was 177 kg and 18.5 taxa; and the mean prediction error (as shown in Fig. 3) for the seven models ranged from 118 kg (random forest hurdle) to 148 kg (random forest with class balancing), and 11 taxa (random forest hurdle) to 31 taxa (JSDM non-spatial). The mean number of taxa observed was 24, compared to predicted which ranged from 39 (GAMM hurdle) to 127 (JSDM non spatial).

Table 1. Comparison of model predictive performance, based on the number of taxa categorized by median AUC (a) and RMAEp (b) from cross-validation. Only hurdle models are included for AUC as they contained a binomial model component. The cumulative number of taxa summed in the direction of decreasing performance appears in parentheses. The model with the best performance, i.e. the highest number of cumulative taxa as we move through the categories, is highlighted in grey; a second model is highlighted if it is within one taxon. RMAEp is the proportional change in RMAE compared to an intercept-only model, e.g. if there are 50 taxa with RMAEp scores > 0.35, this indicates that 50 taxa reduced the error (RMAE) of an intercept-only model by > 35%; this is similar to saying that model explained > 35% of the variation for those taxa. Each model has two rows, one for all 130 taxa (first row) and one for 30 "primary" and "key secondary" species identified in the Fishery Management Plan for the ocean trawl fishery (second row, grey text, Table A4).

(a)	**AUC > 0.9**	**AUC 0.8-0.9**	**AUC 0.7-0.8**	**AUC < 0.7**
GAMM Hurdle	6	17 (23)	28 (51)	79 (130)
	2	*5 (7)*	*10 (17)*	*13 (30)*
RF Hurdle	21	51 (72)	48 (120)	10 (130)
	6	*15 (21)*	*9 (30)*	*0 (30)*
RF Hurdle Down	26	57 (83)	39 (122)	8 (130)
	8	*14 (22)*	*8 (30)*	*0 (30)*
Hmsc non-spatial	22	40 (62)	47 (109)	21 (130)
	7	*12 (19)*	*9 (28)*	*2 (30)*
Hmsc spatial	20	40 (60)	47 (107)	23 (130)
	6	*13 (19)*	*9 (28)*	*2 (30)*

(b)	**RMAEp > 0.35**	**RMAEp 0.25-0.35**	**RMAEp 0.15-0.25**	**RMAEp < 0.15**
GAMM Hurdle	50	33 (83)	18 (101)	29 (130)
	11	*6 (17)*	*7 (24)*	*6 (30)*
GAMM Tweedie	57	32 (89)	25 (114)	16 (130)
	11	9 (20)	4 (24)	6 (30)
RF Hurdle	60	49 (109)	14 (123)	7 (130)
	18	*10 (28)*	*2 (30)*	*0 (30)*
RF Hurdle Down	13	10 (23)	15 (38)	92 (130)
	5	*3 (8)*	*5 (13)*	*17 (30)*
RF	32	43 (75)	39 (114)	16 (130)
	10	*10 (20)*	*9 (29)*	*1 (30)*
Hmsc non-spatial	35	39 (74)	44 (118)	12 (130)
	9	*9 (18)*	*7 (25)*	*5 (30)*
Hmsc spatial	68	45 (113)	15 (128)	2 (130)
	16	*10 (26)*	*2 (28)*	*2 (30)*

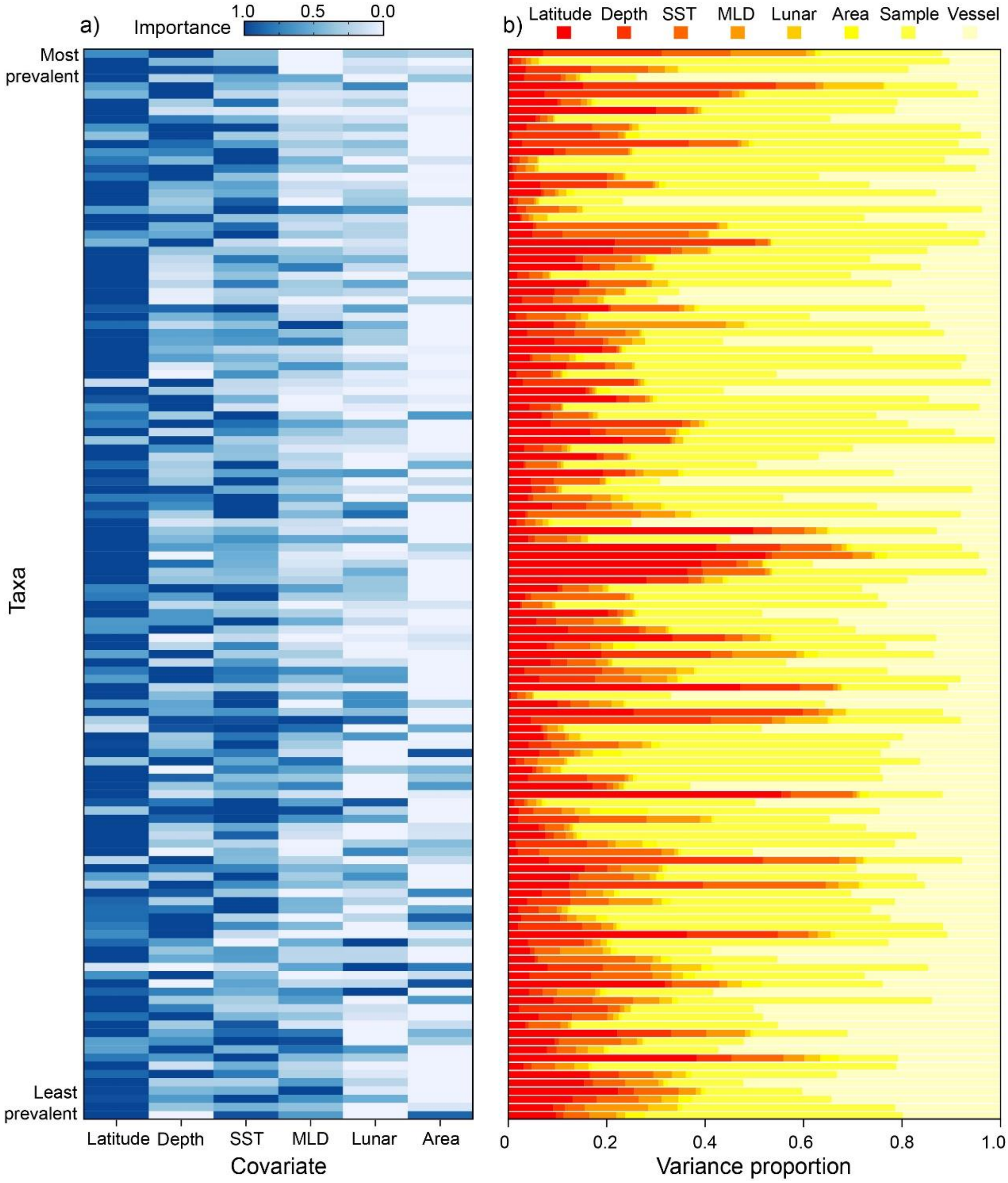


Fig. 1. a) Relative importance of model covariates for all 130 taxa for the presence component of the random forest hurdle model, and b) variance partitioning for all taxa from the presence component of the spatial JSDM. Taxa are ordered in rows from most prevalent (top) to least prevalent (bottom). In a) colours are the importance metric estimated by the randomForest R package but scaled from 0 (least important covariate) to 1 (most important covariate). The large amount of dark blue for 'Latitude' indicates that this covariate had a high, or the highest, importance value for most taxa. Due to our rescaling this plot shows the relative importance of covariates to each other, rather than their importance relative to the data. In b) colours are the variance components for the six covariates plus two random effects (Sample, which is at the observation level and spatially structured; and fishing Vessel). Mean variance proportions across taxa are Latitude=11.1%, Depth=7.5%, SST=6.6%, MLD=3.2%, Lunar=1.7%, Area=1.0%, Sample=42.2%, and Vessel=26.6%.

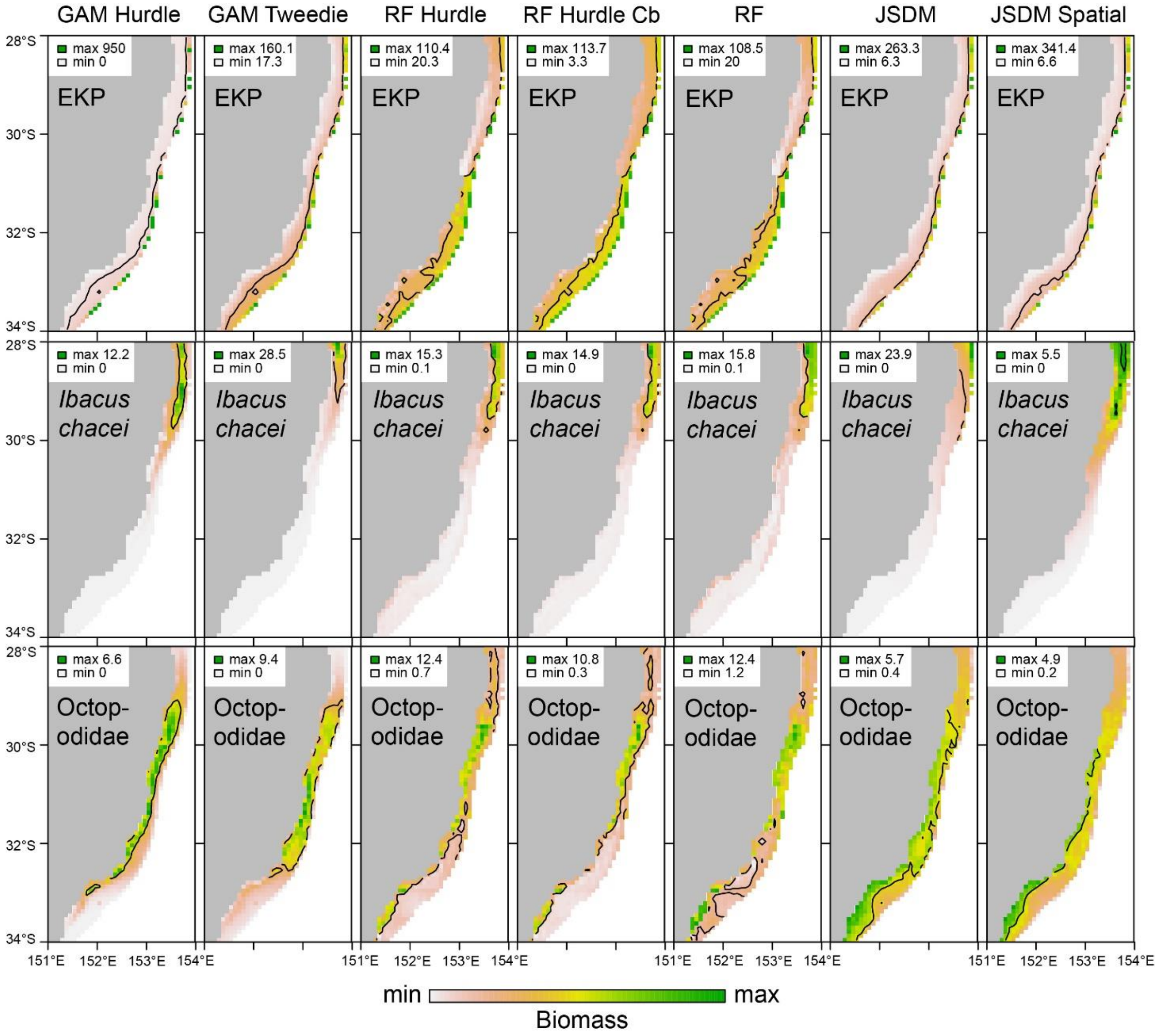


Fig. 2. Maps of predicted biomass of three illustrative taxa (rows), across the seven model types (columns): GAMM hurdle, GAMM Tweedie, random forest (RF) hurdle, RF hurdle with class-balancing (Cb), RF, JSDM, and JSDM with spatial random effects. The three taxa are Eastern king prawn (EKP, *Melicertus plebejus*), smooth (Balmain) bug *(Ibacus chacei*), and Octopodidae. The units for predicted biomass are kg per trawl fishing a constant 0.44 km^2 (the mean effort). Each plot has its own colour bar, but for simplicity a single bar is shown (bottom) and the minimum and maximum units of that colour bar given in each plot (e.g. for GAM hurdle EKP the darkest green indicates 950 kg per trawl and the lightest grey indicates 0 kg per trawl). These maps are for the example date 01-05-2016, and the SST and MLD values used for prediction are from that date; lunar illumination was fixed at 0.5. The prediction region was limited to the fishable depths (< 270 m, 150 fathoms). The black line in each plot is a contour line following an arbitrary but constant value for each taxon, and is used to aid model comparisons.

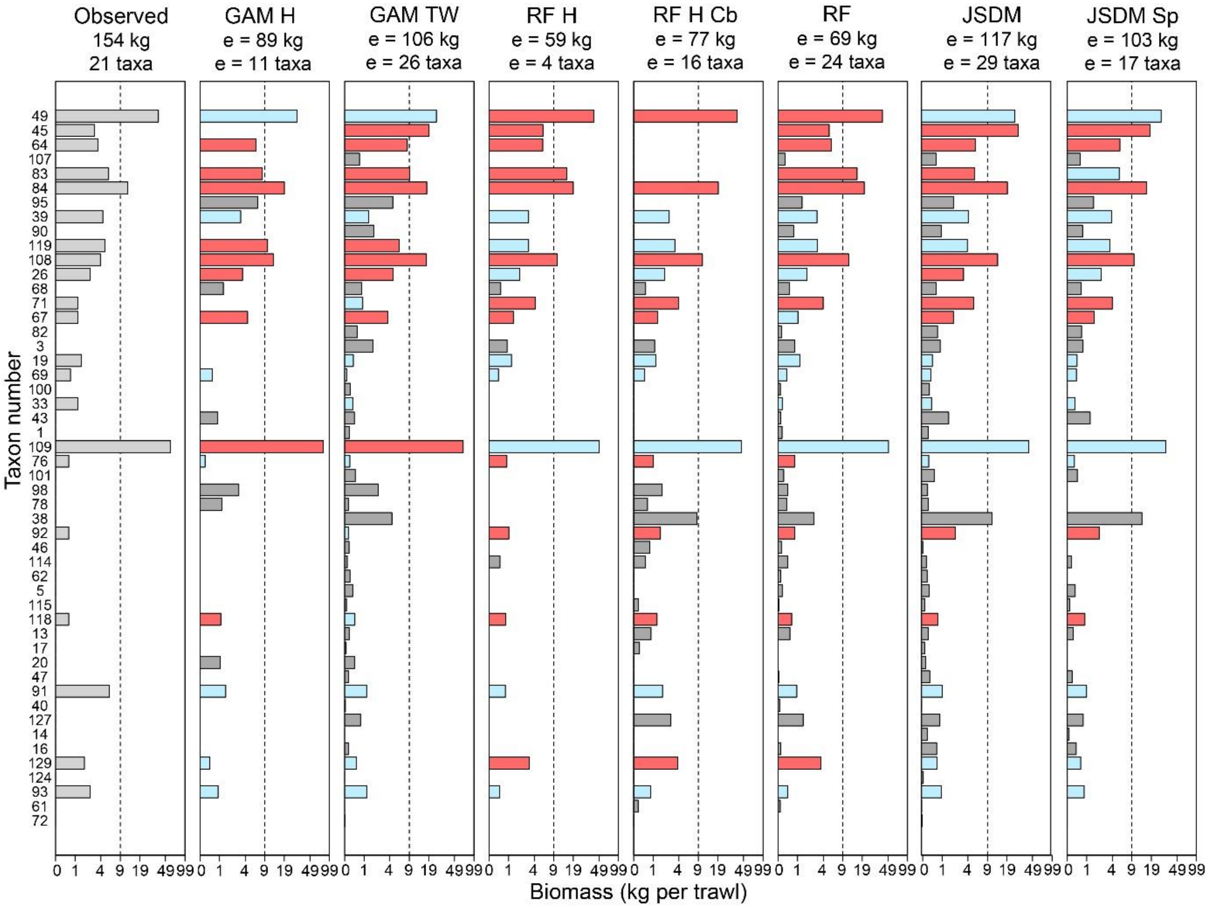


Fig. 3. Barplots comparing observed (left) and predicted catch compositions for the seven model types, for a single example trawl. For the predicted catches, bar colour shows if the predicted biomass for a taxon is higher (red) or lower (blue) than observed, or if the predicted taxon was not present in the observed trawl (dark grey). Only the most prevalent 50 taxa are shown for clarity. The total observed biomass and taxon richness for the top 50 taxa in this trawl was 154 kg and 21 taxa, and the error (e) in these metrics is listed for each model type. For example, the random forest hurdle model (RF H) over- or under-estimated the observed biomass by 59 kg, and 4 taxa that weren’t observed were predicted to occur or vice versa. The x-axes are (biomass +1) log-transformed for clarity of large and small biomasses, and the x-axis labels have been corrected (-1). See Fig. 2 caption for model type abbreviations. Taxon numbers correspond to those in Table A1; some key ones are: 49 *Melicertus plebejus*; 83 *Platycephalus caeruleopunctatus*, 39 *Ibacus chacei*, 109 *Sillago robusta*, 91 *Portunus pelagicus*.

4. Discussion

4.1. Predictive performance

Trawl fisheries often catch many taxa, and we've shown that modelling these fisheries does not necessarily require extensive aggregation of taxa, and correlative multi-species modelling (including joint species modelling) is a viable tool for exploring spatial and temporal aspects of catch composition and its management. Our broad analysis created a useful screening study to evaluate whether it is worthwhile to include multi-species prediction as part of a management strategy. We found that there is potential for using catch prediction as a management tool in our study fishery, with 80% of the taxa across a wide spectrum of prevalence showing at least some useful predictive skill. We also found that prediction error can still be large – a likely pattern for all highly mixed fisheries – so great value will come from more targeted analyses that evaluate predictive performance against metrics tailored for specific management objectives, e.g. threshold numbers of catches per season for a specific group of taxa, or the presence of taxa inside and outside proposed closures. More targeted objectives and applications may also alter how the SDM predictions are calculated, especially whether a study focuses on predicting trawl biomass or 'species richness' (Guillera-Arroita et al 2015, Muscatello et al 2021).

Our choice of model covariates was standard, with catch compositions well explained by constant spatial variables (latitude and depth) and a dynamic variable representing seasonal variation (predominantly SST). Customizing SDMs for each taxon would be challenging, and we approached our analysis here like a repeatable survey program might: using a single set of covariates for all taxa. However, customizing model structure could have value for further predictive value, especially given the diversity of taxa in trawl catches (e.g. teleosts, elasmobranchs, crustaceans, molluscs). JSDMs are more restricted in terms of model structure, so stacked SDMs have an advantage when model structure or covariates are to be varied among taxa.

The JSDMs showed that much of the explained information can be attributed to the random effects, and that much of this has spatial structure (Fig. 1b). We consider it likely this represents some missing covariates, and a key covariate would be bottom habitat and its complexity. However, the spatial scale of the random effect is potentially large (~150 km) so there may also be larger subregional variation. In any case, further evaluation of potential covariates is worthwhile and may increase the predictive performance of catch composition. In general, our study shows that whole-catch compositions of highly mixed fisheries can have predictive potential, but the level of predictive skill will be case-dependent and depend on the data available (Guillera-Arroita et al 2015, Norberg et al 2019). Other considerations are species prevalence and the sampling patterns inherent in fishery-dependent data (Karp et al 2023). The presence of very common taxa was generally predicted with high discrimination. For example, our target species Eastern king prawn (EKP) was almost always present in trawls, but we have little information about conditions when it is absent. Consequently, EKP is predicted to almost always be caught (and has high AUC in the hurdle models), yet the predicted biomass was only of moderate accuracy (Table A4, rank 59). This does not mean we cannot predict poorer EKP habitat, we do (although extrapolation is prone to error; Fig. 2), but this result indicates that we need trawl data outside areas targeted by fishers (likely from an independent survey) to better estimate out-of-sample accuracy and for even moderate levels of extrapolation. This is likely true for all very prevalent taxa.

4.2. Modelling approach

The best model in general was the random forest hurdle, although the random forest hurdle model with class balancing and the spatial JSDM also showed high performance in some cases. While our GAMMs had a moderate wiggliness (due in part here to avoid the number of coefficients exceeding observations for rare species) random forests were more flexible, and could model small-scale spatial hotspots of a taxon. Our cross validation showed this did not lead to overfitting, although other forms of cross validation (Roberts et al 2017) targeted to specific management objectives might show different results. We did not fit complex spatial smoothers, and in such cases GAMMs could be a high performing method. Class balancing showed promise for the random forest model by improving discrimination. Even though class balancing inflated presence and thus biomass, the predicted spatial distributions were realistic and the fitted responses may be more accurate of smoother, and thus more realistic, environmental responses for rare taxa (Stock et al 2020, Smith et al 2021). So, the use of class balancing, while not essential, may have value as a member of a model ensemble tested against more specific objectives.

Our analysis finds that both stacked and joint species modelling frameworks are useful for prediction. Previous research has found that the choice of framework is unimportant compared to other modelling decisions (Zurell et al 2020), although variance can be underestimated in stacked models (Harris 2015). Our analysis shows that a JSDM is not necessarily the best model for prediction, it does provide considerable additional information on ‘community assembly’, which is our case includes how the community interacts with the fishing gear. If a JSDM is used spatial random effects seem essential (Norberg et al 2019). Our particular set of species traits did not appear to add to predictive performance, but taxonomic relatedness showed more skill at explaining catch composition, and including this as an input when jointly estimating environmental drivers may be an advantage of JSDMs. However, we also found that our JSDM was prone to extrapolation outside the regions of the data (Fig. 2), whereby unsampled depths predict higher biomasses than observed. This was likely due to the limitations of the linear and quadratic relationships in the GLM. So, although the spatial JSDM’s performance was good, this would not be the case for predictions much outside the region used to fit the data. Although a model like the JSDM may be highly efficient at fitting data and generating unbiased inference (Ovaskainen et al 2016), this does not necessarily translate to predictive power (Poggiato et al 2021). When predicting to new values, the latent variables are averaged, and this can greatly reduce a JSDM’s impressive goodness-of-fit. An exception to this is conditional prediction, in which joint model predicts some new information in observed samples (Zurell et al 2020). In our case, the comparatively high performance of the spatial JSDM shows that room for innovation may be the prediction of unobserved discards from landed species, i.e. conditional joint prediction (Wilkinson et al 2020, Vallé et al 2023).

4.3. Management relevance

There is already some spatial management of the OPT, with numerous inshore areas closed to benthic trawl nets to protect nursery areas (Taylor et al 2021a), and there is noted potential for adaptive management (Taylor and Johnson 2020, Camp et al 2023). These forms of management are only viable due to the predictability of the spatial-temporal distribution of species and their catches. Currently, this closure network benefits some target (e.g. EKP, stout whiting) and abundant bycatch species (e.g. gurnards, longspine flathead; Taylor et al 2020, Taylor et al 2021b), however further multi-species modelling would inform the utility of spatial management to address the issue of multi-species bycatch (Pons et al 2022). Such models could be used to identify and avoid high-bycatch risk fishing, perhaps through the use of dynamic bycatch risk maps. When this was done post-hoc on real data from a groundfish fishery, bycatch-to-target ratios decreased by around 50% when removing the top 10% of high-risk fishing effort (Stock et al 2020). Well-fitting models across many species are essential for this to be a viable process.

We chose to focus on predictive skill at the level of an individual trawl, but there are other levels which could be useful for key taxa or groups, such as the abundance caught per season or broad management area. These coarser levels benefit from more data, and some taxa which showed poor predictive skill here at a trawl level might show useful skill at a coarser level. Thus, the objective of catch prediction is important to consider when evaluating predictive skill, and some objectives such as predicting annual bycatch rates of threatened, endangered, and protected species (TEPS) may be skilful with very little data (Breiner et al 2015, Jannot et al 2021). In our case, 143 TEPS across seven taxa were caught in 37 trawls (Johnson and Barnes 2023), but only one aggregated taxon (Syngnathidae) was abundant enough to be included in our analysis and it showed only moderate predictive skill at a trawl level (Table A4, rank 97). This means that management tools such as spatial targeting (which require a fine resolution of prediction) are unlikely to be viable for the management of TEPS catch in the OPT due to their rarity, but it doesn’t mean that TEPS catches cannot be predicted at some level, and thus influence management strategies such as closures or effort control. Given that most of the bycatch biomass in the OPT comes from common taxa, and most show reasonable predictive skill (Table A4), spatial-temporal management of bycatch, including minimizing the impact on target catch (Smith et al 2021, Pons et al 2022), remain viable tools worth exploring for this fishery.

Acknowledgements

We are grateful to the fishermen from the New South Wales Ocean Prawn Trawl fishery for their participation in the observer program. We are also grateful to the following people for the observing work or data entry: Matt O’Sullivan, Luke Van Lawick, Andy Schwartz, Kristian Gerathy, Wade Sutherland, Bradley Leach, Tristian New, Mitchell Burns, Alice Pidd and Adam Wiltshire. Sampling was done under permit P01/0059(A)-2.0 and Animal Research Authority NSW DPI 07/03.

Funding

Funding for this work was provided by the New South Wales Commercial Fishing Trust (Project no. RDE581-1). Infrastructure and logistic support were provided by the New South Wales Department of Primary Industries.

Data availability

The observer data used in this analysis is confidential and cannot be shared. To access an aggregated and anonymised version contact the authors. R code for this study is available at https://github.com/smithja16/MultiSpeciesModels_Observer.

References:

Barnes, T. C., Candy, S. G., Morris, S., & Johnson, D. D. (2022). Understanding discarding in trawl fisheries: A model based demersal case study with implications for mitigating and assessing impacts. Plos one, 17(2), e0264055.

Breiman, L. (2001). Random Forests. Machine Learning 45, 5–32.

Breiner, F. T., Guisan, A., Bergamini, A., & Nobis, M. P. (2015). Overcoming limitations of modelling rare species by using ensembles of small models. Methods in Ecology and Evolution, 6(10), 1210-1218.

Broadhurst, M. K. (2000). Modifications to reduce bycatch in prawn trawls: a review and framework for development. Reviews in Fish Biology and Fisheries, 10, 27-60.

Broadhurst, M. K., Brand, C. P., & Kennelly, S. J. (2012). Evolving and devolving bycatch reduction devices in an Australian penaeid-trawl fishery. Fisheries Research, 113(1), 68-75.

Camp, E. V., Johnson, D. D., & Taylor, M. D. (2023). Modelling the potential consequences of adaptive closure management in a penaeid trawl fishery. Aquaculture and Fisheries, 8(2), 190-201.

Catchpole, T. L., Ribeiro-Santos, A., Mangi, S. C., Hedley, C., & Gray, T. S. (2017). The challenges of the landing obligation in EU fisheries. Marine Policy, 82, 76-86.

DAWE (2021). Assessment of the New South Wales Ocean Trawl Fishery. Department of Agriculture, Water and the Environment, Commonwealth of Australia, June 2021. https://www.dcceew.gov.au/sites/default/files/env/pages/37cbb8f7-caee-45dd-9070-d44c43ed9a63/files/nsw-ocean-trawl-assessment-2021.pdf

Dolder, P. J., Thorson, J. T., & Minto, C. (2018). Spatial separation of catches in highly mixed fisheries. Scientific reports, 8(1), 13886.

DPI (2007). Fishery Management Strategy for the NSW Ocean Trawl Fishery. NSW Department of Primary Industries. https://www.dpi.nsw.gov.au/__data/assets/pdf_file/0005/632408/Ocean-Trawl-FMS.pdf

Ferrier, S., & Guisan, A. (2006). Spatial modelling of biodiversity at the community level. Journal of applied ecology, 43(3), 393-404.

Fletcher M. & Fisk, G. (2017). NSW Marine Estate Statewide Threat and Risk Assessment. Final Report for the Marine Estate Management Authority, August 2017. https://www.marine.nsw.gov.au/__data/assets/pdf_file/0011/1352666/NSW-Marine-Estate-Threat-and-Risk-Assessment-Final-Report.pdf

Guillera-Arroita, G., Lahoz-Monfort, J. J., Elith, J., Gordon, A., Kujala, H., Lentini, P. E., ... & Wintle, B. A. (2015). Is my species distribution model fit for purpose? Matching data and models to applications. Global ecology and biogeography, 24(3), 276-292.

Harris, D. J. (2015). Generating realistic assemblages with a joint species distribution model. Methods in Ecology and Evolution, 6(4), 465-473.

Hui, F. K., Warton, D. I., Foster, S. D., & Dunstan, P. K. (2013). To mix or not to mix: comparing the predictive performance of mixture models vs. separate species distribution models. Ecology, 94(9), 1913-1919.

Jannot, J. E., Ward, E. J., Somers, K. A., Feist, B. E., Good, T. P., Lawson, D., & Carretta, J. V. (2021). Using Bayesian models to estimate humpback whale entanglements in the United States west coast sablefish pot fishery. Frontiers in Marine Science, 8, 775187.

Johnsen, J. P., & Eliasen, S. (2011). Solving complex fisheries management problems: What the EU can learn from the Nordic experiences of reduction of discards. Marine Policy, 35(2), 130-139.

Johnson, D. D. & Barnes, T. C. (2023). Observer-based survey of the prawn trawl sectors (inshore & offshore prawn) of the New South Wales ocean trawl fishery. Retained and discarded catch characteristics. Fisheries Final Report Series No. 163. NSW Department of Primary Industries.

Karp, M. A., Brodie, S., Smith, J. A., Richerson, K., Selden, R. L., Liu, O. R., ... & Pozo Buil, M. (2023). Projecting species distributions using fishery-dependent data. Fish and Fisheries, 24(1), 71-92.

Kennelly, S. J. (1995). The issue of bycatch in Australia's demersal trawl fisheries. Reviews in Fish Biology and Fisheries, 5, 213-234.

Kennelly, S. J. (2020). Bycatch Beknown: Methodology for jurisdictional reporting of fisheries discards–Using Australia as a case study. Fish and Fisheries, 21(5), 1046-1066.

Kennelly, S. J., Liggins, G. W., & Broadhurst, M. K. (1998). Retained and discarded by-catch from oceanic prawn trawling in New South Wales, Australia. Fisheries Research, 36(2-3), 217-236.

Liaw A. and Wiener M. (2002). Classification and Regression by randomForest. R News 2(3), 18-22.

Maunder, M. N., & Punt, A. E. (2004). Standardizing catch and effort data: a review of recent approaches. Fisheries research, 70(2-3), 141-159.

Muscatello, A., Elith, J., & Kujala, H. (2021). How decisions about fitting species distribution models affect conservation outcomes. Conservation Biology, 35(4), 1309-1320.

Norberg, A., Abrego, N., Blanchet, F. G., Adler, F. R., Anderson, B. J., Anttila, J., ... & Ovaskainen, O. (2019). A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. Ecological monographs, 89(3), e01370.

Ovaskainen, O., Roy, D. B., Fox, R., & Anderson, B. J. (2016). Uncovering hidden spatial structure in species communities with spatially explicit joint species distribution models. Methods in Ecology and Evolution, 7(4), 428-436.

Ovaskainen, O., Tikhonov, G., Norberg, A., Guillaume Blanchet, F., Duan, L., Dunson, D., ... & Abrego, N. (2017). How to make more out of community data? A conceptual framework and its implementation as models and software. Ecology letters, 20(5), 561-576.

Ovaskainen, O., & Abrego, N. (2020). Joint species distribution modelling: With applications in R. Cambridge University Press.

Panzeri, D., Russo, T., Arneri, E., Carlucci, R., Cossarini, G., Isajlović, I., ... & Libralato, S. (2023). Identifying priority areas for spatial management of mixed fisheries using ensemble of multi-species distribution models. Fish and Fisheries.

Poggiato, G., Münkemüller, T., Bystrova, D., Arbel, J., Clark, J. S., & Thuiller, W. (2021). On the interpretations of joint modeling in community ecology. Trends in Ecology & Evolution, 36(5), 391-401.

Pons, M., Watson, J. T., Ovando, D., Andraka, S., Brodie, S., Domingo, A., ... & Hilborn, R. (2022). Trade-offs between bycatch and target catches in static versus dynamic fishery closures. Proceedings of the National Academy of Sciences, 119(4), e2114508119.

Poos, J. J., Bogaards, J. A., Quirijns, F. J., Gillis, D. M., & Rijnsdorp, A. D. (2010). Individual quotas, fishing effort allocation, and over-quota discarding in mixed fisheries. ICES Journal of Marine Science, 67(2), 323-333.

R Core Team (2023). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Robert, M., Calderwood, J., Radford, Z., Catchpole, T., Reid, D. G., & Pawlowski, L. (2019). Spatial distribution of discards in mixed fisheries: species trade-offs, potential spatial avoidance and national contrasts. Reviews in Fish Biology and Fisheries, 29, 917-934.

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., ... & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography, 40(8), 913-929.

Roberts, S. M., Halpin, P. N., & Clark, J. S. (2022). Jointly modeling marine species to inform the effects of environmental change on an ecological community in the Northwest Atlantic. Scientific Reports, 12(1), 132.

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. and Müller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics, 12, p. 77.

Soto, M., Fernández-Peralta, L., Rey, J., Czerwisnki, I., García-Cancela, R., Llope, M., ... & Pennino, M. G. (2023). Modelling drivers of trawl fisheries discards using Bayesian spatio-temporal models. Fisheries Research, 268, 106830.

Stock, B. C., Ward, E. J., Eguchi, T., Jannot, J. E., Thorson, J. T., Feist, B. E., & Semmens, B. X. (2020). Comparing predictions of fisheries bycatch using multiple spatiotemporal species distribution model frameworks. Canadian Journal of Fisheries and Aquatic Sciences, *77*(1), 146-163.

Smith, J. A., Tommasi, D., Welch, H., Hazen, E. L., Sweeney, J., Brodie, S., ... & Jacox, M. G. (2021). Comparing dynamic and static time-area closures for bycatch mitigation: a management strategy evaluation of a swordfish fishery. Frontiers in Marine Science, 8, 630607.

Taylor, M. D., & Johnson, D. D. (2020). Evaluation of adaptive spatial management in a multi-jurisdictional trawl fishery. Regional Studies in Marine Science, 35, 101206.

Taylor, M. D., Silburn, J., Booth, D. J., & Johnson, D. D. (2020). Impact of spatial management on nontarget species in an oceanic penaeid trawl fishery. North American Journal of Fisheries Management, 40(2), 509-520.

Taylor, M. D., Hale, D., & Johnson, D. D. (2021a). Biological evaluation of an existing spatial management network for an exploited penaeid species. Regional Studies in Marine Science, 47, 101924.

Taylor, M. D., Johnson, D. D., Hale, D., Camp, E. V. (2021b) Spatial management within the NSW Ocean Trawl Fishery. Report 2016/020. Final Report to the Fisheries Research and Development Corporation, NSW Department of Primary Industries, Port Stephens.

Thorson, J. T., Ianelli, J. N., Larsen, E. A., Ries, L., Scheuerell, M. D., Szuwalski, C., & Zipkin, E. F. (2016). Joint dynamic species distribution models: a tool for community ordination and spatio-temporal monitoring. Global Ecology and Biogeography, 25(9), 1144-1158.

Thorson, J. T., & Barnett, L. A. (2017). Comparing estimates of abundance trends and distribution shifts using single-and multispecies models of fishes and biogenic habitat. ICES Journal of Marine Science, 74(5), 1311-1321.

Thorson, J. T., Maureaud, A. A., Frelat, R., Mérigot, B., Bigman, J. S., Friedman, S. T., ... & Wainwright, P. (2023). Identifying direct and indirect associations among traits by merging phylogenetic comparative methods and structural equation models. Methods in Ecology and Evolution, 14(5), 1259-1275.

Tikhonov, G., Opedal, Ø. H., Abrego, N., Lehikoinen, A., de Jonge, M. M., Oksanen, J., & Ovaskainen, O. (2020). Joint species distribution modelling with the R-package Hmsc. Methods in ecology and evolution, 11(3), 442-447.

Vallé, C., Poggiato, G., Thuiller, W., Jiguet, F., Princé, K., & Le Viol, I. (2023). Species associations in joint species distribution models: from missing variables to conditional predictions. Journal of Biogeography.

Wilkinson, D. P., Golding, N., Guillera-Arroita, G., Tingley, R., & McCarthy, M. A. (2021). Defining and evaluating predictions of joint species distribution models. Methods in Ecology and Evolution, 12(3), 394-404.

Wood, S. N. (2017). Generalized additive models: an introduction with R. CRC press.

Zurell, D., Zimmermann, N. E., Gross, H., Baltensweiler, A., Sattler, T., & Wüest, R. O. (2020). Testing species assemblage predictions from stacked and joint species distribution models. Journal of Biogeography, 47(1), 101-113.

Zuur, A.F., Ieno, E.N., Walker, N., Saveliev, A.A., and Smith, G.M. (2009). Mixed effects models and extensions in ecology with R. 1st ed. Springer, New York.

APPENDIX

Table A1. Taxa listed alphabetically with their taxon number used in some plots. Also reported are the two traits and their values specified in their spatial joint species distribution model.

Taxon num.	**Taxon**	**Asymptotic length cm**	**Habitat category**
1	Anoplocapros.inermis	35.7	Demersal
2	Antennariidae.Tetrabrachiidae.Lophichthyidae	35.2	Demersal
3	Aptychotrema.rostrata	130	Benthic
4	Atypichthys.strigatus	18	Reef
5	Bathysauridae.Synodontidae	30.05	Benthic
6	Brachyura	15	Benthic
7	Caesionidae.Lutjanidae.Symphysanodontidae	46.85	Demersal
8	Caproidae	24.1	Demersal
9	Carangidae	43.7	Coastal_pelagic
10	Cardinalfish.spp (Apogonidae)	13.5	Reef
11	Centroberyx.affinis	26	Reef
12	Centropogon.australis	12	Reef
13	Charybdis.spp	15	Benthic
14	Choerodon.spp	28.2	Reef
15	Chrysophrys.auratus	72.3	Demersal
16	Class.Asteroidea	20	Benthic
17	Class.Echinoidea	15	Benthic
18	Class.Gastropoda	15	Benthic
19	Congridae.Colocongridae	75.6	Benthic
20	Crossorhombus.valderostratus	20.99	Benthic
21	Cynoglossidae	27.79	Benthic
22	Dactyloptena.spp	22.88	Benthic
23	Dasyatidae	111.4	Benthic
24	Dinolestes.lewini	55	Reef
25	Diodontidae.spp	35.7	Demersal
26	Draconettidae.Callionymidae	34.02	Benthic
27	Epinephelus.ergastularius	105.8	Reef
28	Fistulariidae	22.7	Demersal
29	Frogfish.spp (Batrachoididae)	30	Reef
30	Genypterus.spp	106.9	Demersal
31	Gerreidae	25	Reef
32	Glaucosoma.scapulare	67.1	Reef
33	Gonorynchus.greyi	40	Benthic
34	Halieutaea.spp	25.3	Benthic
35	Halimochirurgus.spp	25	Demersal
36	Helicolenus.spp	45	Reef
37	Heterodontus.spp	150	Demersal
38	Hypnos.monopterygius	50	Benthic
39	Ibacus.chacei	20	Benthic
40	Ibacus.spp	20	Benthic
41	Jasus.verreauxi	50	Benthic
42	Lactoria.spp	20	Demersal
43	Lagocephalus.cheesemanii	22.2	Demersal
44	Latridae	53.4	Reef
45	Lepidotrigla.spp	24.9	Benthic

Table A1 cont.

Taxon num.	Taxon	Asymptotic length cm	Habitat category
46	Loliginidae	40	Coastal_pelagic
47	Lophonectes.gallus	21	Benthic
48	Majidae	40	Benthic
49	Melicertus.plebejus	18	Benthic
50	Melo.umbilicatus	15	Benthic
51	Metapenaeus.macleayi	13	Benthic
52	Meuschenia.scaber	17.7	Demersal
53	Meuschenia.trachylepis	17.7	Demersal
54	Microcanthus.strigatus	15	Reef
55	Monacanthidae	17.7	Demersal
56	Monocentrididae	24.6	Reef
57	Moridae.spp	58.4	Demersal
58	Mullidae	29.2	Demersal
59	Myliobatidae	108.6	Benthic
60	Nelusetta.ayraudi	50	Demersal
61	Nemipteridae	28.8	Reef
62	Neosebastes.spp.	29.6	Reef
63	Nototodarus.gouldi	40	Coastal_pelagic
64	Octopodidae	40	Benthic
65	Ophichthidae	72.4	Benthic
66	Opistognathidae	15	Benthic
67	Optivus.agastos	10	Reef
68	Order.Stomatopoda	15	Benthic
69	Order.Teuthoidea	50	Coastal_pelagic
70	Orectolobus.spp	162	Demersal
71	Other.flounder (Plueronectiformes)	30	Benthic
72	Paguroidea.spp	10	Benthic
73	Pandalidae	15	Benthic
74	Parapercis.spp	10	Reef
75	Paraplagusia.bilineata	27.8	Benthic
76	Parupeneus.spp	29.7	Demersal
77	Pectinidae	10	Benthic
78	Penaeoidea.Caride	15	Benthic
79	Penaeus.spp	20	Benthic
80	Pentacerotidae	47.9	Demersal
81	Pinguipedidae	35.6	Benthic
82	Platycephalus.conatus	49.3	Benthic
83	Platycephalus.caeruleopunctatus	44.1	Benthic
84	Platycephalus.longispinis	49.3	Benthic
85	Platycephalus.richardsoni	61.4	Benthic
86	Platycephalus.spp	49.3	Benthic
87	Pomacentridae.spp	15.7	Reef
88	Pomatomus.saltatrix	96.6	Coastal_pelagic
89	Porcellanidae.undifferentiated	15	Benthic
90	Portunidae	20	Benthic

Table A1 cont.

Taxon num.	Taxon	Asymptotic length cm	Habitat category
91	Portunus.pelagicus	20	Benthic
92	Portunus.rubromarginatus	20	Benthic
93	Portunus.sanguinolentus	20	Benthic
94	Priacanthidae	36.5	Reef
95	Priacanthus.macracanthus	33.5	Reef
96	Pseudophycis.spp	60.8	Demersal
97	Pseudorhombus.jenynsii	36.3	Benthic
98	Pseudorhombus.tenuirastrum	36.3	Benthic
99	Rajidae	103.5	Benthic
100	Samaris.cristatus	31.8	Benthic
101	Saurida.spp	39.2	Reef
102	Scomber.australasicus	36	Coastal_pelagic
103	Scorpaenidae.spp	30.2	Reef
104	Scyliorhinidae	107.3	Demersal
105	Scyllaridae.spp	25	Benthic
106	Sepia.spp	20	Coastal_pelagic
107	Sepiidae	20	Coastal_pelagic
108	Sillago.flindersi	26.3	Benthic
109	Sillago.robusta	18.5	Benthic
110	Soleidae	27.8	Benthic
111	Squatinidae	138	Benthic
112	Syngnathidae	16.1	Reef
113	Terapontidae	29.1	Reef
114	Tetraodontidae	29.1	Demersal
115	Tetrosomus.reipublicae	30	Demersal
116	Torpedinidae.Narcinidae.Hypnidae	117.67	Benthic
117	Trachichthyidae	35.4	Demersal
118	Trachinocephalus.myops	30	Benthic
119	Trachurus.novaezelandiae	32	Coastal_pelagic
120	Triakidae.spp	139.7	Demersal
121	Trichiurus.lepturus	96.8	Demersal
122	Triglidae.spp	34.1	Benthic
123	Trygonoptera.testacea	45	Benthic
124	Trygonorrhina.spp	130	Benthic
125	Upeneichthys.lineatus	29.2	Demersal
126	Uranoscopidae	41.9	Benthic
127	Urolophidae.spp	80.1	Benthic
128	Volutidae	20	Benthic
129	Zebrias.scalaris	20	Benthic
130	Zeus.faber	54.5	Demersal

Model details

The GAMMs were fitted using the 'mgcv' R package (Wood 2017). The hurdle model consisted of a logit and gamma distributions. The single model used a Tweedie distribution. Each covariate was fitted using a thin-plate regression spline, with maximum wiggliness controlled using the k parameter (k=5 in most cases). The vessel random effect and MLD were removed from the abundance-only component of the hurdle model, due to avoid the number of coefficients exceeding the number of observations for the rarest species. For the GAMMs, an unexpected issue for very common and abundant species was the gamma model occasionally leading to very large predicted catches (and so large to become negative on the inverse link scale), which indicates poor performance of this model and distribution. In these cases we restricted the predictions to the maximum observed catch for that species. We also evaluated a log link rather than inverse link and the final predictions were very similar. R code to fit all models can be found at https://github.com/smithja16/MultiSpeciesModels_Observer.

The random forests were fitted using the 'randomForest' R package (Liaw and Wiener 2002). The hurdle model used classification trees and regression trees (Stock et al 2020). The single model used regression trees. All models were fit with two variables randomly sampled at each split (mtry=2) and 1200 trees, based on tuning of key species models. In classification tasks it is important to avoid mismatches in the proportions of the minority and majority classes (here presences and absences), because this can lead to misleading estimates of accuracy (e.g. a rarely observed taxa may appear to be predicted well by a model that classifies everything as zeros). For this class balancing, we used downsampling or synthetic minority oversampling (SMOTE) to ensure similar class proportions (Stock et al 2020). Downsampling works by randomly reducing the majority class observations (e.g. trawls with zero catch of a taxa) to match the number of minority class observations (e.g. trawls with catches of that taxa). SMOTE works by simultaneously downsampling the majority class as well as oversampling the minority class (Chawla et al 2002), which it does by creating 'synthetic' observations with explanatory values derived from linear combinations of n=5 real minority class observations. Because SMOTE creates synthetic observations, we restricted SMOTE to the rarest or most common taxa. Testing showed this was best used for taxa with fewer than 140 instances of a minority class (~66 taxa) which is around 10% prevalence; downsampling was used for all other taxa. Downsampling was done using the 'caret' R package (Kuhn 2008), and SMOTE using the 'smotefamily' R package (Siriseriwan 2019).

The latent variable JSDMs were fitted using the 'Hmsc' R package (Tikhonov et al 2020a). This was done only as a hurdle model, using probit and normal distributions, with the response variable log-transformed for the model with the normal distribution. We fitted two variants, one with species associations implemented as non-spatial latent variables, and one with taxonomic traits, a phylogenetic tree, and species associations implemented as spatial latent variables. We selected these two variants because the latter is very complex and slow to fit, and the difference between variants allows us to evaluate whether the added complexity is required. Spatial latent variables allow residual information to be spatially structured, and taxonomic traits and a phylogeny allow us to evaluate the extent that similar responses to the environment relate to similar traits or relatedness. The spatial latent variables (i.e. random effects) can improve predictions by borrowing information across both species and location (Ovaskainen and Abrego 2020). We used the Gaussian predictive process method (which uses knots) to fit the spatial latent variables, as an approximation for big spatial data (Tikhonov et al 2020b).

Two traits were included for each taxon in the spatial JSDM models, asymptotic length (cm) and a general habitat classification (benthic, demersal, reef, coastal pelagic); these are listed in Table A1. For fish taxa, asymptotic lengths were taken from the R package FishLife (Thorson et al 2020, Thorson 2023). FishLife is based on information from FishBase (fishbase.org) but uses models to predict life-history parameter values, which means that phylogenetic trait imputation is used to predict values for species without information (or for broader taxonomic levels) from related (or included) species. In a few cases, asymptotic lengths were taken from local information when FishLife was inaccurate. Asymptotic lengths for invertebrates were approximated using local knowledge and information from SeaLifeBase (sealifebase.org). Habitat categories were based on expert opinion and existing classifications (Truong et al 2017, Goddard et al 2022). The phylogenetic tree for the final suite of species/taxa (n=130) was calculated using taxonomic relatedness (i.e. a taxonomic tree), thus specifying phylogenetic distance = 1 for each level of the tree (Thorson et al 2023), and was built using the ‘ape’ R package (Paradis and Schliep 2019). Compared to a phylogenetic tree, a taxonomic tree will likely overestimate the distance between closely related taxa, and vice versa for distantly related taxa, but there was insufficient alternate information available.

In the JSDM the SST, MLD, and lunar covariates were fitted as quadratic terms, and all covariates were centred and scaled to encourage model convergence. Hmsc uses Bayesian MCMC, and our posterior estimates were derived from three chains, a thinning interval of 20, and 30,000 iterations excluding the first 10,000 as burn-in. We used Hmsc’s default priors, and as recommended (because this uses a normal distribution) we centred and scaled the response variable of the abundance-only component. We evaluated model convergence by evaluating trace plots, effective sample size, and scale reduction factors (Ovaskainen and Abrego 2020). As a tip to other users, Hmsc models are complex and slow to fit and a key element of speeding them up in our study was replacing R’s default algebra libraries with optimised ones. For this we used OpenBLAS (github.com/xianyi/OpenBLAS), which increased speeds by a factor of 2-3. Also see the Hmsc vignette on performance for more information. Even so, the fitting and cross-validation for the spatial model was computationally expensive and increases greatly above ~1000 samples and ~100 taxa. Also, the number of polynomial terms and the number of knots in a spatial grid can both increase run times. Our advice is to start small and build up. Also, running models in parallel was not always faster (sometimes slower) so use that feature carefully. We recommend using a PC with at least 32 GB RAM to avoid memory issues.

Models predicting probability of presences often use an adjustment or threshold to ensure realistic numbers of taxa are predicted per observation unit (e.g. per trawl). We applied thresholding for this reason to all GAMM and random forest models. For the hurdle models, which calculate a presence-absence component, we used the minimum training presence threshold (Guillera-Arroita et al 2015) which uses each taxon's prevalence as the threshold (number of presences:number of samples; to a maximum of 0.95) below which a taxon is considered absent. Our approach was to use this as a truncated threshold which is a more robust approach (Muscatello et al 2021). This truncation means that above the threshold we retain the probability of presences, rather than assuming all probability values greater than the threshold are equal to 1. For the non-hurdle models, we used a minimum biomass threshold, again to ensure some true zeros were predicted and not just very small values. We used half the minimum observed biomass for each taxon as that taxon's biomass threshold, which ensured we were not too strict when truncating predictions. We admit there are many approaches to adjusting and thresholding (Liu et al 2005, Guillera-Arroita et al 2015), and we chose one that created a good balance between discrimination and accuracy, while avoiding a discretization of the output which can lose information (Calabrese et al 2014, Guillera-Arroita et al 2015). Thresholding was not used for the JSDMs, because the biomass estimates were attained by taking the median of many realizations, each of which multiplied together one posterior sample from the presence and abundance-only components where the presence component generated a binary 0-1 value.

We tested including 'Trip ID' as an additional random effect, to explore whether accounting for residual correlation in trawl catches caught during the same trip improved model predictions. This was not added by default due to the unacceptable fit time of the spatial JSDM, and the inability to include this term in the random forests. We tested this random effect in the GAMM framework and found that predictive performance (as in Table 1) was about the same or decreased slightly, so we feel that excluding this covariate did not bias our results.

Appendix References:

Calabrese, J. M., Certain, G., Kraan, C., & Dormann, C. F. (2014). Stacking species distribution models and adjusting bias by linking them to macroecological models. Global Ecology and Biogeography, 23(1), 99-112.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. Journal of artificial intelligence research, 16, 321-357.

Goddard, B. K., Becker, A., Harasti, D., Smith, J. A., Subramaniam, R. C., & Suthers, I. M. (2022). The trophic basis of fish assemblages in temperate estuarine and coastal ecosystems. Marine Biology, 169(2), 19.

Guillera-Arroita, G., Lahoz-Monfort, J. J., Elith, J., Gordon, A., Kujala, H., Lentini, P. E., ... & Wintle, B. A. (2015). Is my species distribution model fit for purpose? Matching data and models to applications. Global ecology and biogeography, 24(3), 276-292.

Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. Journal of Statistical Software, 28(5), 1–26. https://doi.org/10.18637/jss.v028.i05

Liaw A. and Wiener M. (2002). Classification and Regression by randomForest. R News 2(3), 18-22.

Liu, C., Berry, P. M., Dawson, T. P., & Pearson, R. G. (2005). Selecting thresholds of occurrence in the prediction of species distributions. Ecography, 28(3), 385-393.

Muscatello, A., Elith, J., & Kujala, H. (2021). How decisions about fitting species distribution models affect conservation outcomes. Conservation Biology, 35(4), 1309-1320.

Ovaskainen, O., & Abrego, N. (2020). Joint species distribution modelling: With applications in R. Cambridge University Press.

Paradis, E., Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics, 35, 526-528.

Siriseriwan W. (2019). smotefamily: A Collection of Oversampling Techniques for Class Imbalance Problem Based on SMOTE. R package version 1.3.1, https://CRAN.R-project.org/package=smotefamily.

Stock, B. C., Ward, E. J., Eguchi, T., Jannot, J. E., Thorson, J. T., Feist, B. E., & Semmens, B. X. (2020). Comparing predictions of fisheries bycatch using multiple spatiotemporal species distribution model frameworks. Canadian Journal of Fisheries and Aquatic Sciences, 77(1), 146-163.

Thorson, J. T. (2020). Predicting recruitment density dependence and intrinsic growth rate for all fishes worldwide using a data-integrated life-history model. Fish and Fisheries, 21(2), 237-251.

Thorson, J. T., (2023). FishLife: Predict Life History Parameters For Any Fish. R package version 3.0.0. http://github.com/James-Thorson-NOAA/FishLife

Tikhonov, G., Opedal, Ø. H., Abrego, N., Lehikoinen, A., de Jonge, M. M., Oksanen, J., & Ovaskainen, O. (2020a). Joint species distribution modelling with the R-package Hmsc. Methods in ecology and evolution, 11(3), 442-447.

Tikhonov, G., Duan, L., Abrego, N., Newell, G., White, M., Dunson, D., & Ovaskainen, O. (2020b). Computationally efficient joint species distribution modeling of big spatial data. Ecology, 101(2), e02929.

Truong, L., Suthers, I. M., Cruz, D. O., & Smith, J. A. (2017). Plankton supports the majority of fish biomass on temperate rocky reefs. Marine biology, 164, 1-12.

Wood, S.N. (2017). Generalized Additive Models: An Introduction with R (2nd edition). Chapman and Hall/CRC.

Table A2. Model covariates and their sources. The spatial JSDM also used longitude to develop the spatial mesh. Longitude was sourced the same was as latitude. [1]https://marine.copernicus.eu (product: GLOBAL_MULTIYEAR_PHY_001_030). [2]https://www.gebco.net. [3]https://cran.r-project.org/web/packages/lunar/.

Covariate	Description	Source for fitting	Source for prediction
Latitude	degrees; the location of trawl start	Recorded by observers using GPS	Latitude of prediction data
Sea surface temperature	°C; seawater potential temperature near surface; daily mean on day of trawl	CMEMS GLORYS12V1 reanalysis product[1]	CMEMS GLORYS12V1 reanalysis product[1]
Mixed layer depth	m; ocean mixed layer thickness defined by sigma theta; daily mean on day of trawl	CMEMS GLORYS12V1 reanalysis product[1]	CMEMS GLORYS12V1 reanalysis product[1]
Depth	ftm; bottom depth at trawl location (approximate mean)	Recorded by observers from vessel sounder (high agreement with bathymetry chart)	GEBCO_2023 bathymetry chart[2]
Lunar illumination	Proportion; lunar illumination on day of trawl; 0=new moon, 1=full moon	'lunar' R package[3]	'lunar' R package on dates in prediction data; or fixed at 0.5 for example maps
Area fished	km^2; total area swept per trawl, calculated as trawl distance × (0.8 × headrope distance)	Headrope distance (m) and distance (km) were recorded by observers, the latter using multiple GPS coordinates per trawl	Input as a constant mean value
Vessel	vessel identifier; included as a random factor in GAMM and JSDM, and a factor in the random forest	Recorded by observers	Excluded from prediction

Table A3. Summary of covariate importance among models. Numbers are the number of taxa for which each covariate was influential (thus the maximum value is 130). Covariate influence is measured: for GAMMs when $p < 0.05$ for that covariate; for random forests when a covariate had the highest, or within 50% of the highest, ‘importance’ score (based on mean decrease in OOB accuracy); and for the JSDMs as posterior support values at 95% support. Hurdle models are split into their presence (P) and abundance-only (A) components. The quadratic terms (^2) were used only in some JSDM models. Blank cells indicate the covariate was excluded from that model. Due to their different calculations, it is unwise to compare the number of taxa among the three main model types (GAMMs, RFs, JSDMs). The goal here is to compare the relative importance of covariates within models. The two covariates with the most influenced taxa are highlighted for each row in two shades of grey.

	Latitude	**Depth**	**SST**	**SST^2**	**MLD**	**MLD^2**	**Lunar**	**Lunar^2**	**Area**
GAMM hurdle P	70	97	87		54		42		21
GAMM hurdle A	80	68	62				49		44
GAMM Tweedie	74	93	92		65		61		50
RF	111	97	114		85		69		61
RF Hurdle P	127	106	119		89		74		50
RF Hurdle A	94	89	96		56		55		49
RF Hurdle Cb. P	126	106	102		86		77		63
JSDM P	79	87	77	68	55	34	40	23	42
JSDM A	65	57	49	41			27	22	19
JSDM Spatial P	77	83	77	66	60	36	36	23	22
JSDM Spatial A	65	57	49	41			27	22	19

Table A4. Taxa ranked by their approximate predictive skill (highest to lowest). For simplicity we present here only results from the random forest hurdle model. Each taxon's rank is determined by the sum of their two ranks among taxa for AUC score and RMAEp (higher values indicate better performance). Prevalence (Prev.) is the proportion of observed trawls in which a taxon occurred. Highlighted in grey are the 26 taxa which make up 80% of discards by biomass, although some of these taxa are frequently also retained. Highlighted with red text are 30 taxa that are, or contain, "primary species" or "key secondary species" identified in the ocean trawl fishery's Fishery Management Plan relevant to the OPT fishery.

Taxa	**Prev.**	**AUC**	**RMAEp**	**Rank**
Sillago.robusta	0.29	0.95	0.61	1
Helicolenus.spp	0.04	0.93	0.58	2
Sepia.spp	0.02	1.00	0.47	3
Centropogon.australis	0.03	0.93	0.52	4
Pandalidae	0.11	0.94	0.48	5
Ibacus.spp	0.18	0.91	0.55	6
Portunus.rubromarginatus	0.25	0.93	0.48	7
Ibacus.chacei	0.64	0.93	0.48	8
Trygonoptera.testacea	0.09	0.93	0.47	9
Heterodontus.spp	0.11	0.90	0.49	10
Pomacentridae.spp	0.03	0.89	0.52	11
Crossorhombus.valderostratus	0.21	0.91	0.46	12
Platycephalus.richardsoni	0.11	0.90	0.45	13
Choerodon.spp	0.17	0.90	0.44	14
Pseudorhombus.tenuirastrum	0.27	0.92	0.42	15
Priacanthidae	0.07	0.95	0.40	16
Paraplagusia.bilineata	0.05	0.87	0.55	17
Terapontidae	0.03	0.89	0.46	18
Penaeus.spp	0.06	0.90	0.44	19
Lophonectes.gallus	0.20	0.88	0.47	20
Lactoria.spp	0.03	0.91	0.41	21
Latridae	0.07	0.85	0.50	22
Pectinidae	0.05	0.89	0.43	23
Meuschenia.scaber	0.15	0.89	0.42	24
Sillago.flindersi	0.51	0.88	0.43	25
Cardinalfish.spp (Apogonidae)	0.04	0.90	0.40	26
Other.flounder (Plueronectiformes)	0.48	0.86	0.46	27
Bathysauridae.Synodontidae	0.23	0.88	0.41	28
Opistognathidae	0.07	0.96	0.34	29
Meuschenia.trachylepis	0.05	0.88	0.40	30
Samaris.cristatus	0.37	0.89	0.39	31
Trichiurus.lepturus	0.06	0.96	0.32	32
Caesionidae.Lutjanidae.Symphysanodontidae	0.03	0.83	0.44	33
Trygonorrhina.spp	0.16	0.85	0.42	34
Loliginidae	0.25	0.87	0.39	35
Pseudophycis.spp	0.02	0.94	0.32	36
Nelusetta.ayraudi	0.10	0.85	0.40	37
Frogfish.spp (Batrachoididae)	0.05	0.87	0.38	38
Jasus.verreauxi	0.03	0.87	0.37	39
Trachinocephalus.myops	0.22	0.86	0.38	40
Rajidae	0.10	0.90	0.32	41
Ophichthidae	0.02	0.85	0.39	42

901 **Table A4 cont.**

Taxa	**Prev.**	**AUC**	**RMAEp**	**Rank**
Penaeoidea.Caridea	0.26	0.84	0.40	43
Metapenaeus.macleayi	0.04	0.88	0.33	44
Anoplocapros.inermis	0.33	0.81	0.42	45
Upeneichthys.lineatus	0.14	0.83	0.39	46
Portunus.sanguinolentus	0.16	0.86	0.36	47
Genypterus.spp	0.10	0.88	0.33	48
Neosebastes.spp.	0.23	0.83	0.38	49
Nototodarus.gouldi	0.11	0.86	0.33	50
Urolophidae.spp	0.17	0.83	0.38	51
Torpedinidae.Narcinidae.Hypnidae	0.03	0.90	0.29	52
Monacanthidae	0.13	0.75	0.51	53
Pseudorhombus.jenynsii	0.13	0.86	0.32	54
Optivus.agastos	0.48	0.75	0.47	55
Scyliorhinidae	0.09	0.78	0.40	56
Nemipteridae	0.16	0.83	0.34	57
Centroberyx.affinis	0.10	0.87	0.30	58
Melicertus.plebejus	0.99	0.98	0.22	59
Halieutaea.spp	0.03	0.83	0.36	60
Halimochirurgus.spp	0.11	0.89	0.28	61
Charybdis.spp	0.22	0.82	0.37	62
Portunus.pelagicus	0.20	0.82	0.37	63
Octopodidae	0.85	0.76	0.42	64
Class.Echinoidea	0.21	0.80	0.38	65
Caproidae	0.09	0.77	0.41	66
Tetraodontidae	0.24	0.74	0.43	67
Saurida.spp	0.28	0.79	0.37	68
Pomatomus.saltatrix	0.07	0.82	0.33	69
Microcanthus.strigatus	0.04	0.83	0.32	70
Soleidae	0.07	0.84	0.30	71
Gerreidae	0.04	0.78	0.38	72
Zebrias.scalaris	0.17	0.77	0.39	73
Platycephalus.conatus	0.46	0.77	0.38	74
Parupeneus.spp	0.29	0.80	0.33	75
Zeus.faber	0.08	0.85	0.29	76
Atypichthys.strigatus	0.02	0.85	0.28	77
Sepiidae	0.74	0.81	0.32	78
Tetrosomus.reipublicae	0.22	0.80	0.32	79
Trachurus.novaezelandiae	0.53	0.77	0.37	80
Carangidae	0.06	0.82	0.29	81
Parapercis.spp	0.06	0.84	0.26	82
Lagocephalus.cheesemanii	0.34	0.79	0.32	83
Platycephalus.longispinis	0.70	0.80	0.29	84
Order.Teuthoidea	0.37	0.81	0.29	85
Brachyura	0.07	0.71	0.40	86
Antennariidae.Tetrabrachiidae.Lophichthyidae	0.12	0.79	0.30	87
Pinguipedidae	0.03	0.79	0.29	88
Portunidae	0.58	0.74	0.35	89

902

Table A4 cont.

Taxa	Prev.	AUC	RMAEp	Rank
Hypnos.monopterygius	0.25	0.77	0.32	90
Draconettidae.Callionymidae	0.49	0.75	0.33	91
Majidae	0.07	0.81	0.26	92
Paguroidea.spp	0.15	0.79	0.28	93
Mullidae	0.07	0.79	0.28	94
Class.Gastropoda	0.06	0.79	0.28	95
Myliobatidae	0.03	0.84	0.15	96
Syngnathidae	0.02	0.75	0.30	97
Pentacerotidae	0.07	0.75	0.31	98
Epinephelus.ergastularius	0.08	0.78	0.27	99
Dactyloptena.spp	0.04	0.78	0.26	100
Gonorynchus.greyi	0.35	0.78	0.25	101
Class.Asteroidea	0.17	0.74	0.29	102
Moridae	0.04	0.79	0.22	103
Priacanthus.macracanthus	0.66	0.74	0.29	104
Cynoglossidae	0.07	0.77	0.26	105
Scorpaenidae	0.06	0.76	0.26	106
Aptychotrema.rostrata	0.43	0.74	0.28	107
Order.Stomatopoda	0.48	0.70	0.30	108
Fistulariidae	0.05	0.69	0.30	109
Dasyatidae	0.05	0.75	0.25	110
Glaucosoma.scapulare	0.13	0.77	0.21	111
Triglidae.spp	0.02	0.78	0.14	112
Uranoscopidae	0.13	0.72	0.27	113
Platycephalus.spp	0.09	0.76	0.21	114
Trachichthyidae	0.07	0.74	0.24	115
Congridae.Colocongridae	0.39	0.73	0.25	116
Triakidae	0.08	0.71	0.26	117
Platycephalus.caeruleopunctatus	0.72	0.74	0.23	118
Diodontidae	0.10	0.73	0.23	119
Squatinidae	0.02	0.72	0.23	120
Chrysophrys.auratus	0.03	0.68	0.25	121
Dinolestes.lewini	0.07	0.73	0.19	122
Lepidotrigla.spp	0.86	0.70	0.22	123
Scyllaridae.spp	0.02	0.67	0.23	124
Scomber.australasicus	0.06	0.72	0.05	125
Monocentrididae	0.04	0.68	0.20	126
Orectolobus.spp	0.04	0.69	0.02	127
Volutidae	0.02	0.62	0.15	128
Melo.umbilicatus	0.02	0.66	-0.02	129
Porcellanidae	0.04	0.59	0.00	130

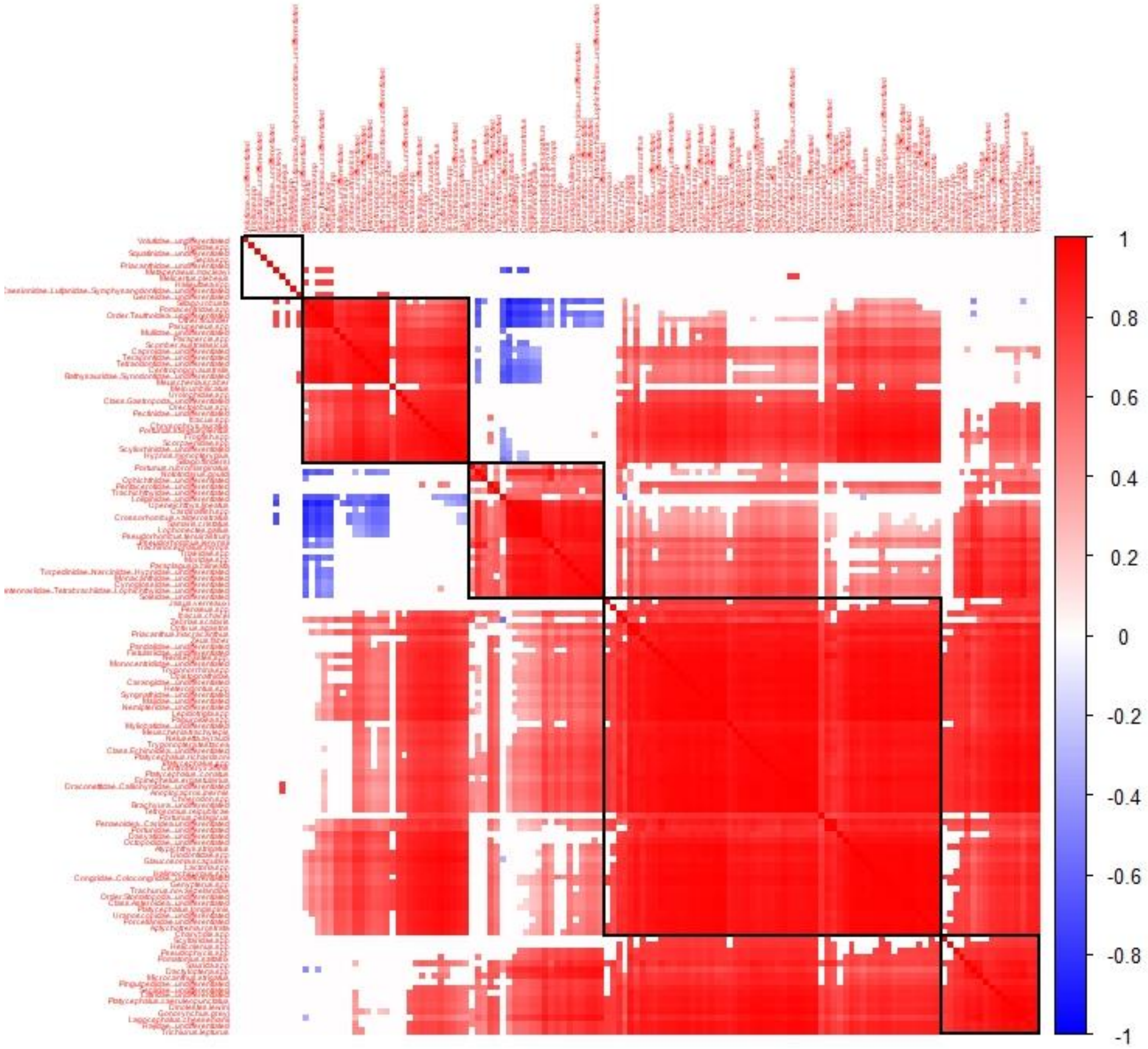


Fig. A1. Residual correlations at the observation-level from the presence component of non-spatial JSDM, at the 95% support level. Red indicates species that are caught together more often than expected from the model, and blues less often than expected. Strong correlations can indicate a missing covariate, and this matrix supports that. Clusters of taxa are indicated by black squares, and are clustered together based on similarity of their residual correlations. Clusters were made using the ‘Wards D2’ method in the ‘corrplot’ R package (Wei and Simko 2021, https://github.com/taiyun/corrplot), and using a visual selection of a parsimonious number of clusters.

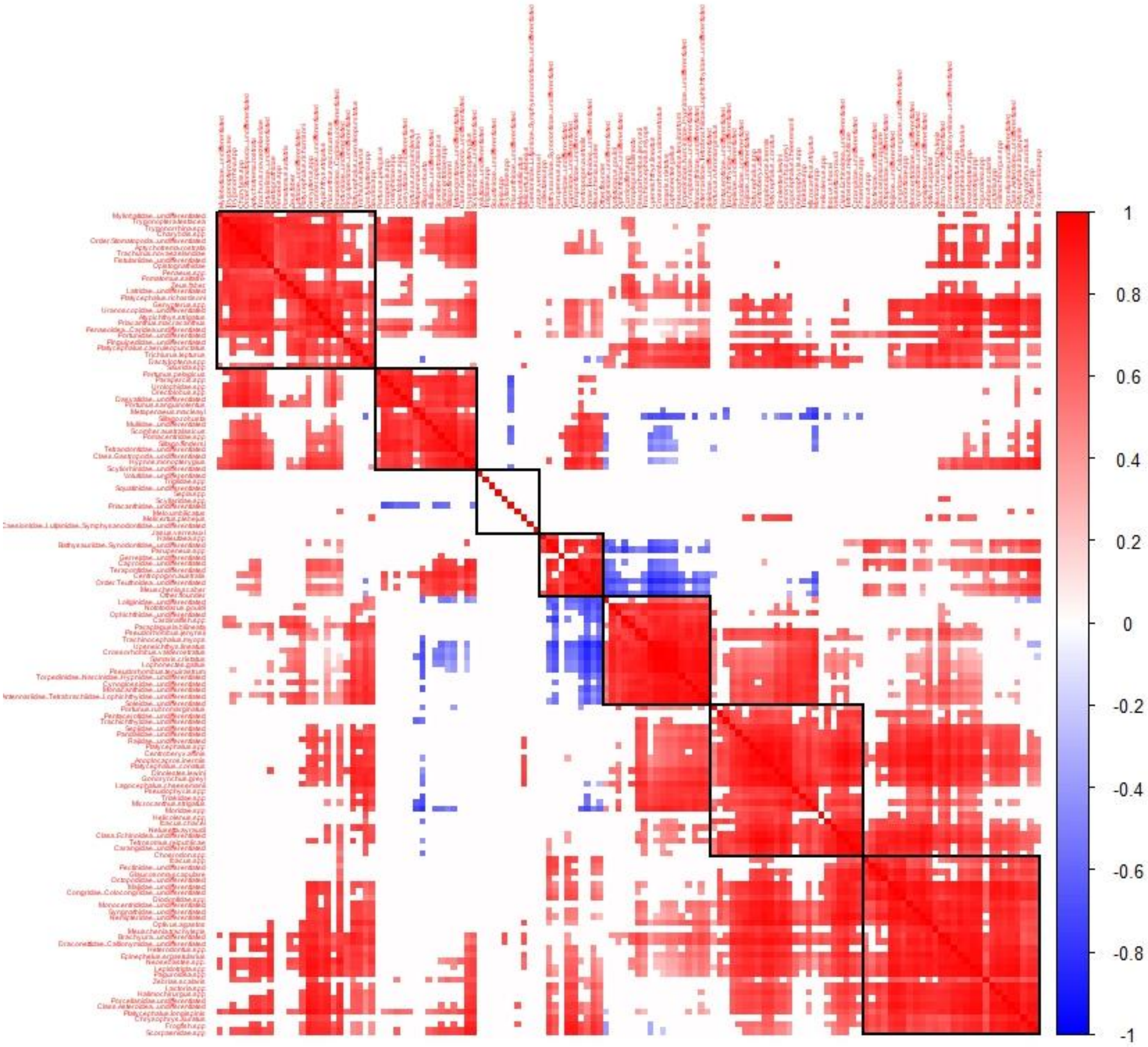
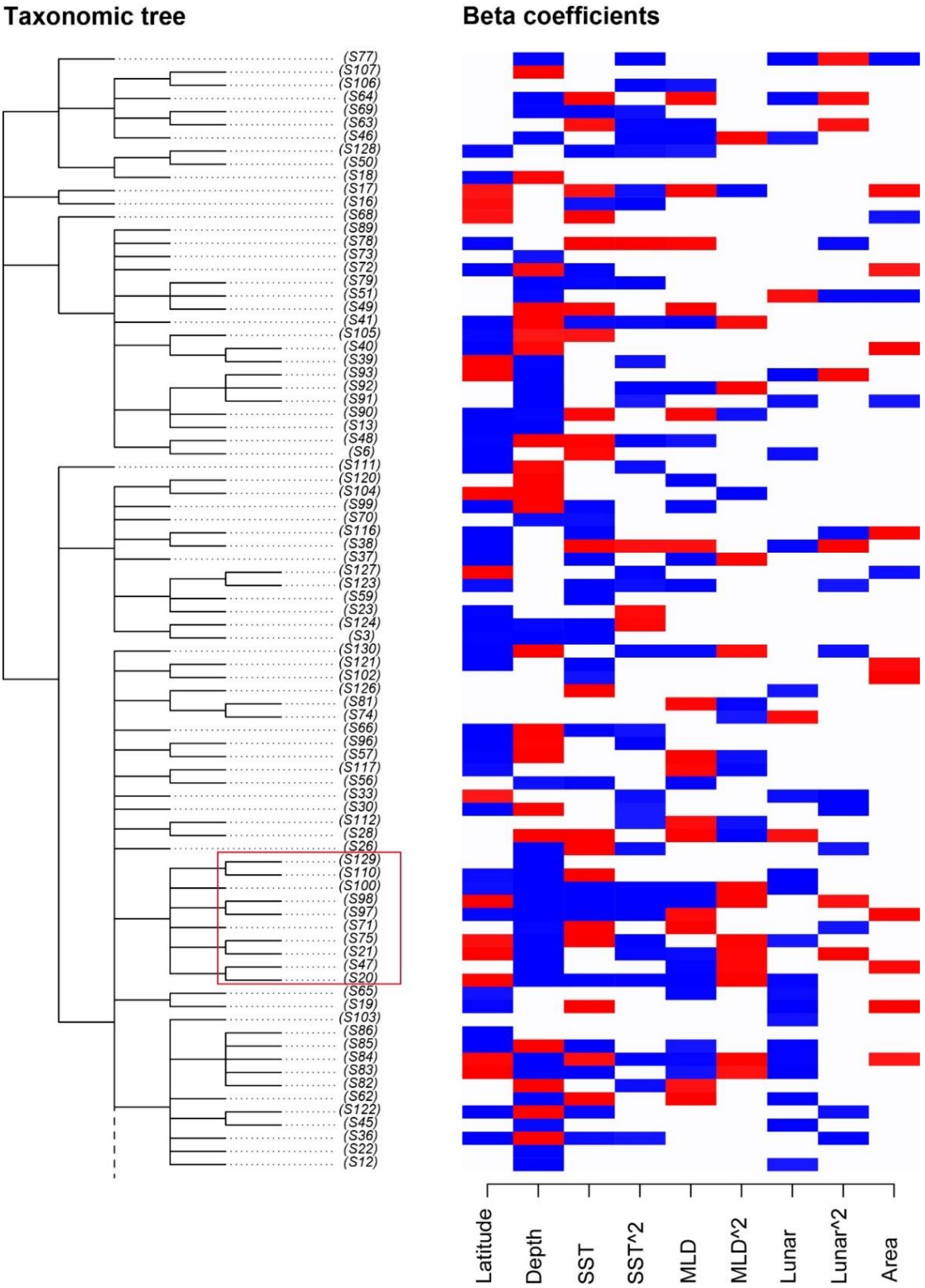


Fig. A2. Residual correlations at the observation-level from the presence component of spatial JSDM, at the 95% support level. Red indicates species that are caught together more often than expected from the model, and blues less often than expected. Strong correlations can indicate a missing covariate, and this matrix has fewer correlations than the non-spatial model (Fig. A1); this is likely due to the spatial random effects in the spatial JSDM explaining much of the residual correlation in the non-spatial JSDM. Clusters of taxa are indicated by black squares, and are clustered together based on similarity of their residual correlations. Clusters were made as in Fig. A1.


Taxonomic tree

Beta coefficients

Fig. A3. Signs of fitted beta coefficients and related taxonomic tree for the spatial JSDM model. Coefficients important with 95% support are coloured red (positive beta coefficient) or blue (negative beta coefficient). The Pleuronectiformes group is indicated by the red box – note that all taxa in this group have a negative beta for the Depth covariate. The plot has been truncated for clarity (lower taxa not shown).

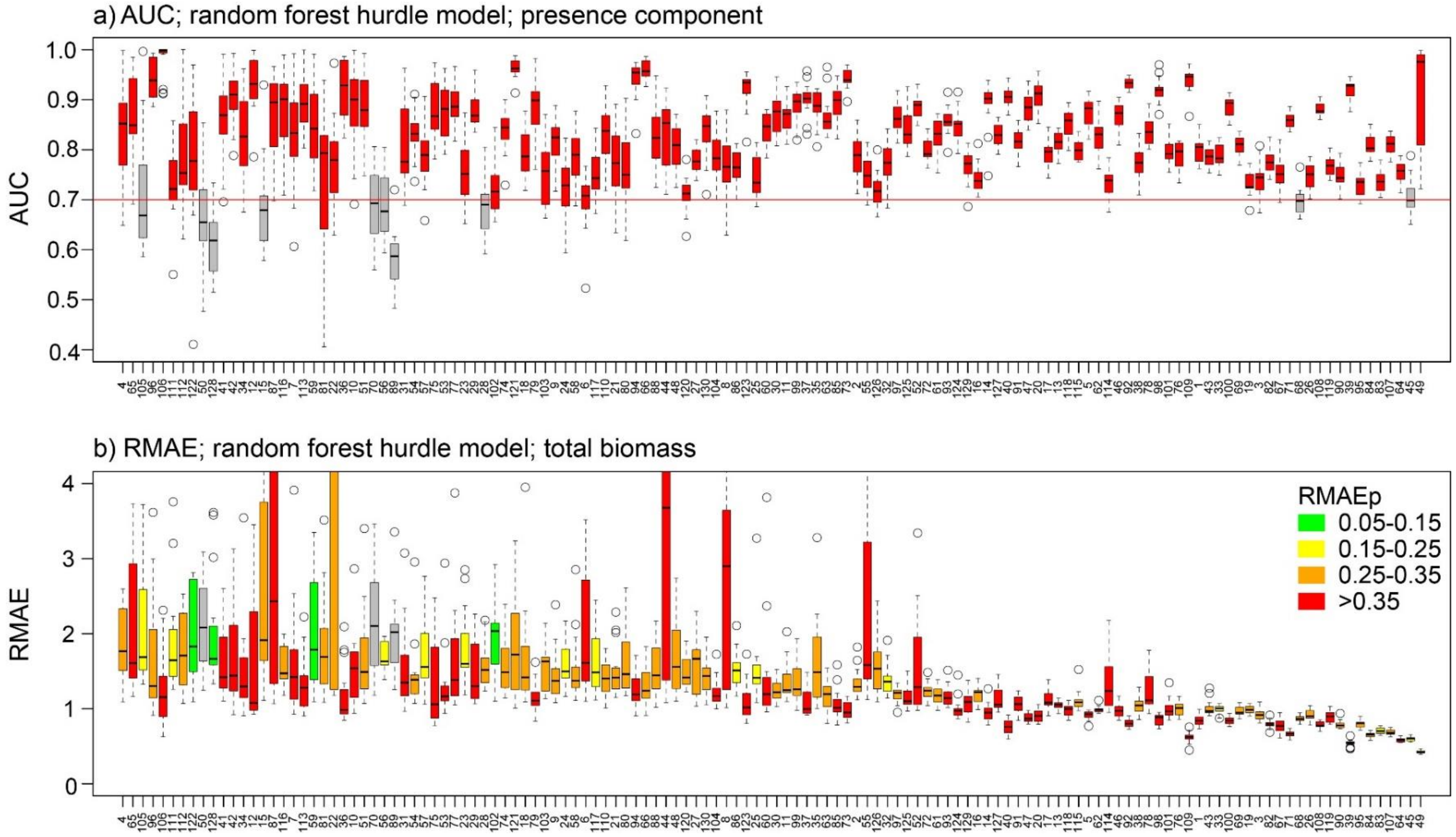


Fig. A4. a) AUC and b) RMAE results from cross-validation for the random forest hurdle model. These shows the spread of results from the 15 folds for each taxon. Taxa are ordered from least prevalent (left) to most prevalent (right), and taxa numbers are listed in Table A1. In a) taxa which have a median AUC > 0.7 are highlighted in red. In b) boxplots are coloured by the respective RMAEp value (the metric summarised in Table 1), with taxa showing the most reduction in RMAE compared to an intercept-only model highlighted in red (> 0.35 proportional reduction in RMAE), and the least improvement in green (0.05-0.15). Although the lowest RMAE values are seen for the most common taxa, the highest proportional improvement is often for the rarer taxa.