Text S1. Statistical considerations for "Distributions of threatened skates and commercial fisheries inform conservation hotspots"

Model description

The general statistical model implemented in the staRVe package is a hierarchically specified spatio-temporal generalized linear mixed model. The hierarchy consists of three levels:

1. the observation equation giving the distribution of the response variable, with observations conditionally independent given a set of spatio-temporal random effects,
2. the spatio-temporal random effects encoding the spatial dependence, which follow a nearest-neighbour Gaussian process (Datta et al. 2016) conditional on a set of purely temporal random effects, and
3. the temporal random effects featuring an AR(1) dependence structure.

Mathematically the model is formulated as

$$\begin{aligned}
\mathbf{Y}_t(\mathbf{s}) &\sim f_\theta\left(y_t(\mathbf{s});\ \mu_t(\mathbf{s}) = g^{-1}\big(\mathbf{X\beta} + \mathbf{w}_t(\mathbf{s})\big)\right) \\
\mathbf{W}_t(\mathbf{s}) &\sim \mathrm{NNGP}\big(\phi \cdot (\mathbf{w}_{t-1}(\mathbf{s}) - \epsilon_{t-1}) + \epsilon_t,\ \mathcal{C}(\mathbf{s}_1, \mathbf{s}_2)\big) \\
\boldsymbol{\epsilon} &\sim \mathrm{N}\left(\mu,\ \frac{\sigma^2}{1-\phi^2} \cdot \Sigma\right) \text{ where } \Sigma_{ij} = \phi^{||i-j||}.
\end{aligned}$$

where t and $\mathbf{s}$ represent time and location respectively, $\mathbf{Y}$ is the observed data, $\mathbf{X}$ is a set of covariates, $\mathbf{W}$ is the set of spatio-temporal random effects, and $\boldsymbol{\epsilon}$ is the set of purely temporal random effects. The parameters to be estimated consist of the parameters θ of the chosen response distribution f, the regression coefficients $\boldsymbol{\beta}$, the spatial covariance matrix $\mathcal{C}$ as a function of the micro-ergodic spatial variance τ, spatial range ρ, and spatial smoothness ν in the Matérn covariance, and finally the temporal variance σ^2, autoregressive parameter ϕ, and mean μ.

The model is implemented in Template Model Builder (TMB) (Kristensen et al. 2016) which applies the Laplace approximation to marginalize over the random effects, and the parameters are estimated by maximum likelihood using a Newton optimizer. Standard errors are computed using automatic differentiation and the delta method. Predictions for the random effects are taken to be the conditional mode of the random effects given the data under the maximum likelihood parameter estimates. More details can be found in (Lawler, Field, and Flemming 2021).

Below we detail a number of technical but important issues for statistical analyses involving spatial regression models. All spatial models based on Gaussian random effects (which is to say, nearly all spatial models) have these issues, so while our model runs into these problems it is far from unique in doing so. We hope that the reader does not come away from reading this with the impression that spatio-temporal generalized linear mixed models should be avoided, but rather that despite their flaws they are still remarkably useful and resilient for producing spatial predictions.

A discussion of the spatial range parameter

The Matérn covariance function has a fairly well-known quirk in that only one of the three parameters needed to describe the function can be consistently estimated. The smoothness parameter ν is typically held at a fixed value with two common choices being the exponential covariance at $\nu = 0.5$ and the Gaussian covariance which is the limit as $\nu \to \infty$. The exponential covariance has the functional form $f(x) = \sigma^2 \text{ Exp}\left[-\frac{x}{\rho}\right]$ where x is the distance between two points, σ^2 is the marginal variance and ρ is the spatial range. Neither the marginal variance nor the spatial parameter can be consistently estimated but the micro-ergodic spatial variance $\tau = \sigma^2/\rho$ can be consistently estimated (Zhang 2004). If one holds either σ^2 or ρ fixed to a known value, then the other can be consistently estimated in the sense that the estimate gives the “right" value of $\hat{\tau}$ given the fixed and known parameter.

While the spatial range cannot be consistently estimated the maximum likelihood estimator often comes close to the true value. For practical purposes the inconsistent estimation of the spatial range does not present much of a challenge since two statistical models can be equivalent, in a sense, even if they have different values of the range parameter, and give identical kriging predictors for prediction at new locations (Zhang 2004).

The performance of models, as measured by predictive performance and inferential ability for τ, using different ML estimation methods for the spatial range was explored in (Kaufman and Shaby 2013). They found that the best method to use is to freely estimate the spatial range parameter jointly with τ. However fixing the spatial range to some relatively large value can work well in practice, particularly when the true spatial range is not too small and the true smoothness is equal to 0.5.

Spatial confounding

A less discussed problem than the estimability of the spatial range parameter is spatial confounding. Initially reported as a discrepancy in regression coefficient estimates between a non-spatial regression and a spatial regression, spatial confounding refers to non-identifiability of regression coefficients for spatially structured covariates (Reich, Hodges, and Zadnik 2006). (Hodges and Reich 2010) introduced restricted spatial regression so that the regression coefficient estimates in a spatial regression matched those of the corresponding non-spatial regression, and the problem now is typically described as the spatial covariate and the unobserved spatial random effects being correlated. Recent studies have shown that non-spatial regression can have high error rates and poor inference for regression coefficients when spatial confounding is present, and unfortunately restricted spatial regression always performs worse than the non-spatial regression (Hanks et al. 2015; Khan and Calder 2019). Both (Hanks et al. 2015), (Page et al. 2017), and (Chiou, Yang, and Chen 2019) each present more or less the same post-hoc correction to restricted spatial regression estimates that gives unbiased and variance-stabilized estimators for the true regression coefficients, however these adjustments require some knowledge of the relationship between unobserved random effects and the spatially varying covariates and are computationally expensive to compute due to the inversion of a large covariance matrix.

The bias (variance) of the regression coefficients estimators in a unrestricted spatial regression is with some assumptions on the data generating process for the covariates directly proportional to the true values of the marginal standard deviation (marginal variance) of the random effects, directly proportional to the correlation (squared correlation) between the covariates and the unobserved random effect, and inversely proportional to the marginal standard deviation (marginal variance) of the covariates (Paciorek 2010; Chiou, Yang, and

Chen 2019). In a seemingly unrelated analysis (Hanks et al. 2015) showed that for spatial covariates and random effects which are independent of each other:

1. as the true spatial range of either quantity tends towards 0, their squared correlation coefficient also tends towards 0, and
2. as the true spatial range of both quantities tend towards infinity, their squared correlation coefficient tends towards 1.

In other words, if both exhibit long-range spatial correlation then the covariate and random effects are quite likely to be collinear and thus exhibit spatial confounding.

Clearly the parameters of the covariance function for the random effects are direct causes of the spatial confounding problem, and the issue is magnified when it comes to jointly estimating the regression coefficients and the spatial covariance parameters. Of course there is the problem that the spatial random effects are unobserved, which always makes inference harder, however this is not the main issue. Even if we had direct observations of the random effects and could treat them as another known covariate in the model, we would still need to account for collinearity between the original covariates and these now known random effects. To de-bias the regression coefficient estimates for the original covariates we would estimate the spatial covariance function for the known random effects and plug-in the estimates to the adjustment given by (Hanks et al. 2015) and (Chiou, Yang, and Chen 2019). If the original covariates and the random effects are positively (negatively) correlated then we would subtract (add) a quantity that is directly proportional to the marginal variance of the random effects: a large marginal variance of the random effects leads to a large negative adjustment to the regression coefficient estimate. However as we noted above the marginal variance is not consistently estimable and two parameter sets for the random effects, one with small marginal variance and one with large marginal variance, can provide identical inference as long as the spatial range is correpondingly small or large to keep their ratio constant. Even if with observations of the random effects, then, the poor estimation of the spatial covariance parameters denies us the ability to correct or de-bias the regression coefficient estimates for the covariates of interest. To our knowledge the connection between spatial confounding and the inconsistent estimation of the covariance parameters has not yet been reported in the literature.

All this to say, once you start including spatially structured covariates to explain variability in a spatially structured response variable your analysis is at the whims of poorly behaved estimators and it can be impossible to say how close you actually are to the right answer based on the data alone. Fortunately, though, while the estimators for the regression coefficients are poorly behaved, it seems as though spatial prediction at new locations is not affected by the spatial confounding problem (or is even slightly better in the presence of spatial confounding) (Page et al. 2017).

For analyses whose goal is to produce smooth maps of a target variable by spatial prediction, parameter estimates, though unreliable, need not be of real concern. Clearly, extrapolation outside the spatial range of the data set will fall victim to the typical mis-specified model problem that all extrapolation methods suffer from. However for spatial prediction within the study area the main output of the analysis will still be statistically sound, even in the absence of correct parameter estimates.

Model selection for spatiotemporal generalized linear mixed models

Model selection for linear models typically consists of the choosing which covariates to include, whether to include random effects, choice of link function, and choice of response distribution using techniques such as comparing information criteria, inspecting residuals, and cross-validation.

The Akaike information criterion (AIC) is a ubiqitous measure, balancing predictive ability with model complexity, used to decide which single configuration of covariates and random effects gives rise to the best model, all else being equal. In mixed models both the marginal AIC which takes into account complexity due to the random effects conditional AIC, which only penalizes the parameters governing the random effects, have been proposed as useful measures for model selection. (Greven and Kneib 2010) argue that both measures are biased where the marginal AIC tends to select models that are too small without random effects and the conditional AIC always selects the model with the most random effects. They present a previously known correction to the conditional AIC for the Gaussian that produces an unbiased AIC, and later provide a method for correcting the conditional AIC for response distributions in the exponential family (Saefken et al. 2014). Their method only works for some members of the exponential family such as the Gaussian, Poisson, and exponential distributions but not for others such as the gamma, negative binomial, and binomial distributions.

Residuals

Residuals for temporal, spatial, and spatio-temporal models need to be defined carefully in order for the residuals to be independent from each other. Take for example a random walk with drift: the naive Pearson residuals defined by $Y_i - \widehat{Y}_i$ will be exhibit temporal dependence even if computed under the correct model (Thygesen et al. 2017). One definition of residuals which results in independence is to take the prediction errors $Y_i - \widehat{Y}_{i\,|i-1}$ where $\widehat{Y}_{i\,|i-1}$ is the predicted value of Y_i given all of the previous data points. Response distributions other than the Gaussian the prediction errors can be checked using randomized quantile prediction residuals. The prediction errors require fitting a number of models equal to the number of data points, to avoid unintentionally borrowing information from "future" data points through the random effects, which can quickly become computationally prohibitive. Residual analysis for spatio-temporal model checking is briefly discussed in (Wikle, Zammit-Mangion, and Cressie 2019) but again these residuals require fitting a model multiple times.

Randomized quantile residuals based on simulations present a more computationally feasible method for model checking since they do not require multiple model fits. However to the best of our knowledge their use for checking spatial or spatio-temporal has not been studied in the literature. While the problems surrounding the spatial range parameter and spatial confounding seem to be somewhat unsolvable, spatio-temporal model validation using simulations presents an area for future research that is clearly needed.

Supplementary References

Chiou YH, Yang HD, Chen CS (2019) An Adjusted Parameter Estimation for Spatial Regression with Spatial Confounding. *Stochastic Environmental Research and Risk Assessment* 33(8): 1535–51. https://doi.org/10.1007/s00477-019-01716-9.

Datta A, Banerjee S, Finley AO, Gelfand AE (2016) Hierarchical Nearest-Neighbor Gaussian Process Models for Large Geostatistical Datasets. *Journal of the American Statistical Association* 111 (514): 800–812. https://doi.org/10.1080/01621459.2015.1044091.

Greven, Sonja, and Thomas Kneib. 2010. "On the Behaviour of Marginal and Conditional Aic in Linear Mixed Models." *Biometrika* 97 (4): 773–89. https://doi.org/10.1093/biomet/asq042.

Hanks EM, Schliep EM, Hooten MB, Hoeting JA (2015) Restricted Spatial Regression in Practice: Geostatistical Models, Confounding, and Robustness Under Model Misspecification. *Environmetrics* 26 (4): 243–54. https://doi.org/10.1002/env.2331

Hodges JS, Reich BJ (2010) Adding Spatially-Correlated Errors Can Mess up the Fixed Effect You Love. *The American Statistician* 64 (4): 325–34. https://doi.org/10.1198/tast.2010.10052

Kaufman CG, Shaby BA (2013) The role of the range parameter for estimation and prediction in geostatistics. *Biometrika* 100 (2): 473–84. https://doi.org/10.1093/biomet/ass079

Khan K, Calder CA (2019) Restricted Spatial Regression Methods: Implications for Inference." *Journal of the American Statistical Association* 2020:1–13. https://doi.org/10.1080/01621459.2020.1788949

Kristensen K, Nielsen A, Berg CW, Skaug HJ, Bell BM (2016) TMB: Automatic Differentiation and Laplace Approximation. *Journal of Statistical Software* 70 (5): 1–21. https://doi.org/10.18637/jss.v070.i05

Lawler E, Field C, Mills Flemming J (2021) Species Distribution Modelling with Spatio-Temporal Nearest Neighbour Gaussian Processes. *arXiv* [stat.ME] (2105.06902). http://arxiv.org/abs/2105.06902.

Paciorek CJ (2010) The Importance of Scale for Spatial-Confounding Bias and Precision of Spatial Regression Estimators. *Statistical Science : A Review Journal of the Institute of Mathematical Statistics* 25 (1): 107–25. https://dx.doi.org/10.1214%2F10-STS326

Page GL, Liu Y, He Z, Sun D (2017) Estimation and Prediction in the Presence of Spatial Confounding for Spatial Linear Models." *Scand J Statist* 44 (3): 780–97. https://doi.org/10.1111/sjos.12275.

Reich BJ, Hodges JS, Zadnik V (2006) Effects of Residual Smoothing on the Posterior of the Fixed Effects in Disease-Mapping Models. *Biometrics* 62 (4): 1197–1206. https://doi.org/10.1111/j.1541-0420.2006.00617.x

Saefken B, Kneib T, van Waveren CS, Greven S (2014) A Unifying Approach to the Estimation of the Conditional Akaike Information in Generalized Linear Mixed Models. *Electronic Journal of Statistics* 8 (1): 201–25. https://doi.org/10.1214/14-EJS881.

Thygesen UH, Albertsen CM, Berg CW, Kristensen K, Nielsen A (2017) Validation of ecological state space models using the Laplace approximation. *Environmental and Ecological Statistics* 24(2):317-339. https://doi.org/10.1007/s10651-017-0372-4

Wikle CK, Zammit-Mangion A, Cressie NAC (2019) *Spatio-Temporal Statistics with R*. Chapman & Hall/Crc the R Series. CRC Press, Taylor & Francis Group. https://books.google.ca/books?id=OdhOswEACAAJ.

Zhang H (2004) Inconsistent Estimation and Asymptotically Equal Interpolations in Model-Based Geostatistics. *Journal of the American Statistical Association* 99 (465): 250–61. https://doi.org/10.1198/016214504000000241.

Table S1. ar1 correlation parameters. Estimated time effect parameters for presence and abundance generalized linear mixed models

Species	Parameter estimate	Standard error	Model
Winter skate	0.9128192	0.05082436	Presence
Thorny skate	0.9950559	0.03789686	Presence
Smooth skate	0.9841288	0.05005849	Presence
Cod	0.958652	0.03831958	Presence
Haddock	1	9.67E-07	Presence
Pollock	0.9917045	0.04168401	Presence
Atlantic halibut	1	0	Presence
Redfish	0.9662924	0.04111907	Presence
Silver hake	1	0	Presence
American plaice	1	0	Presence
Yellowtail flounder	1	0	Presence
Witch flounder	1	1.77E-07	Presence
Winter flounder	1	0	Presence
Winter skate	0.5298323	0.1534382	CPUE
Thorny skate	0.8920145	0.1078976	CPUE
Smooth skate	0.2131272	0.4017496	CPUE
Cod	0.9898238	0.04081878	CPUE
Haddock	0.9233384	0.03095771	CPUE
Pollock	0.8680234	0.06393095	CPUE
Atlantic halibut	1	2.41E-09	CPUE
Redfish	0.9890275	0.02740027	CPUE
Silver hake	0.9499352	0.0353312	CPUE
American plaice	0.9907277	0.03310303	CPUE
Yellowtail flounder	0.9431188	0.03147937	CPUE
Witch flounder	0.9999999	0	CPUE
Winter flounder	0.9686549	0.04309326	CPUE

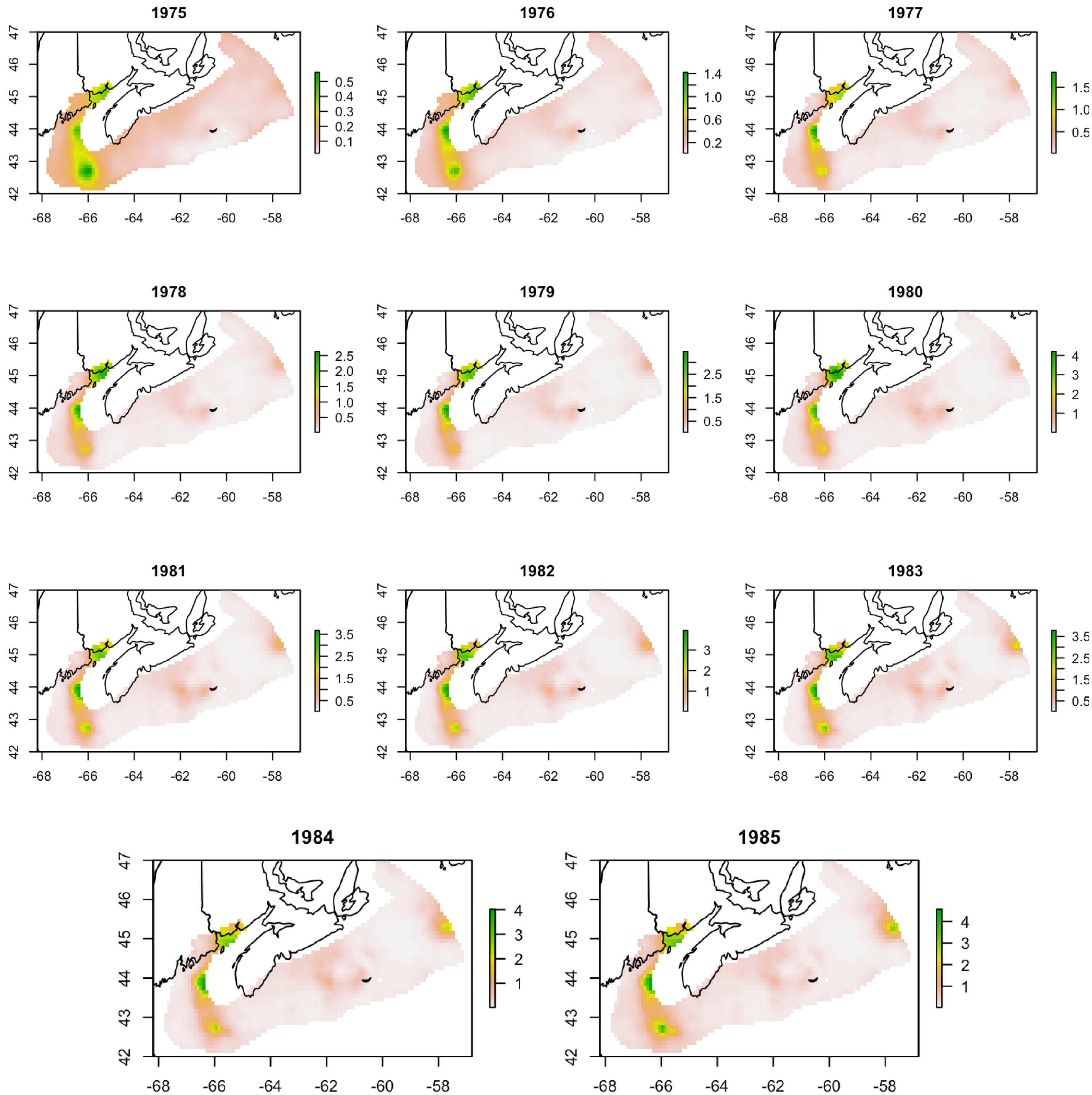

Fig. S1. Standard error of predicted historical thorny skate density. Shown is the calculated standard error for thorny skate density, 1975–1985

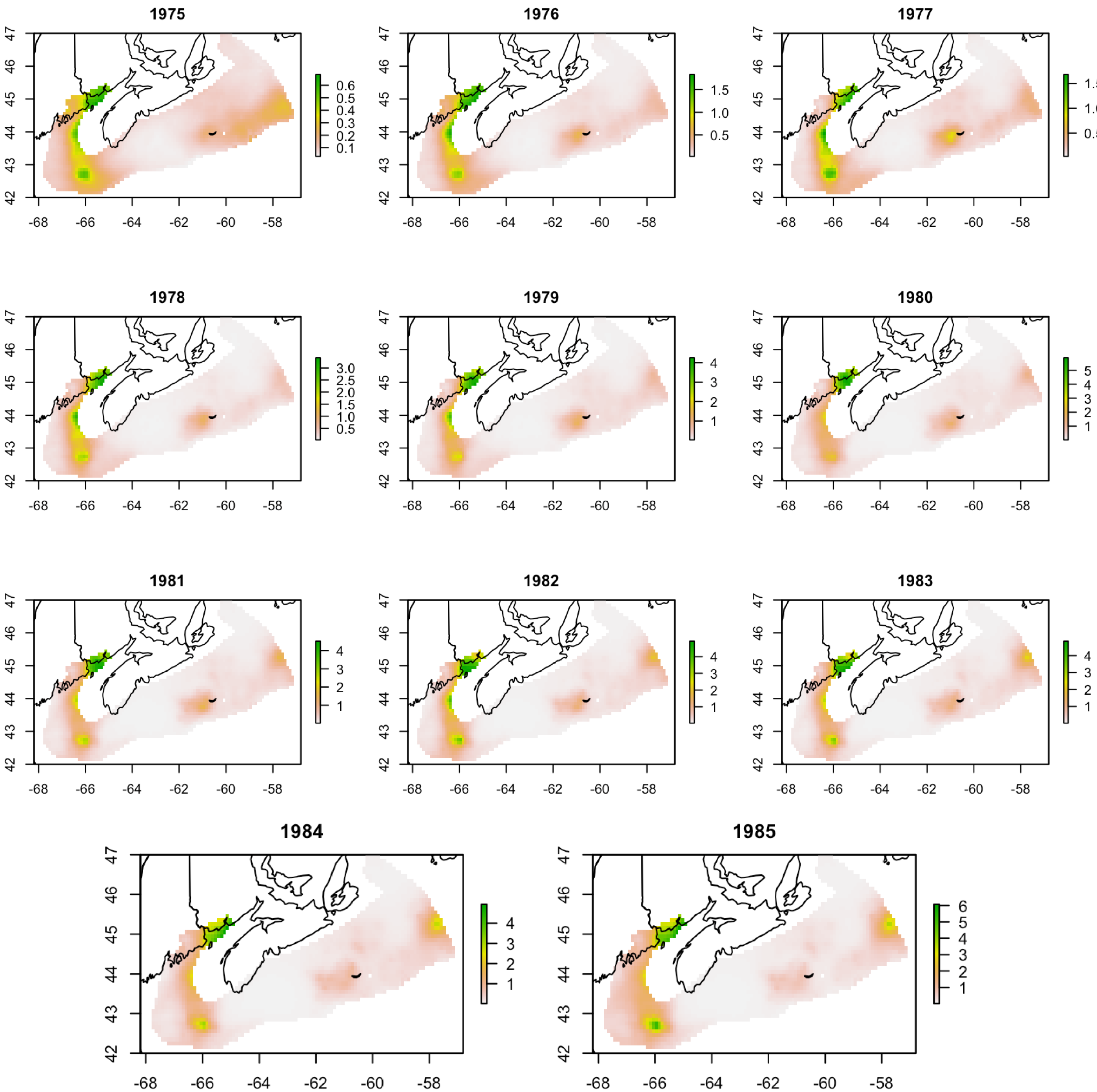

Fig. S2. Standard error of predicted historical winter skate density. Shown is the calculated standard error for winter skate density, 1975–1985

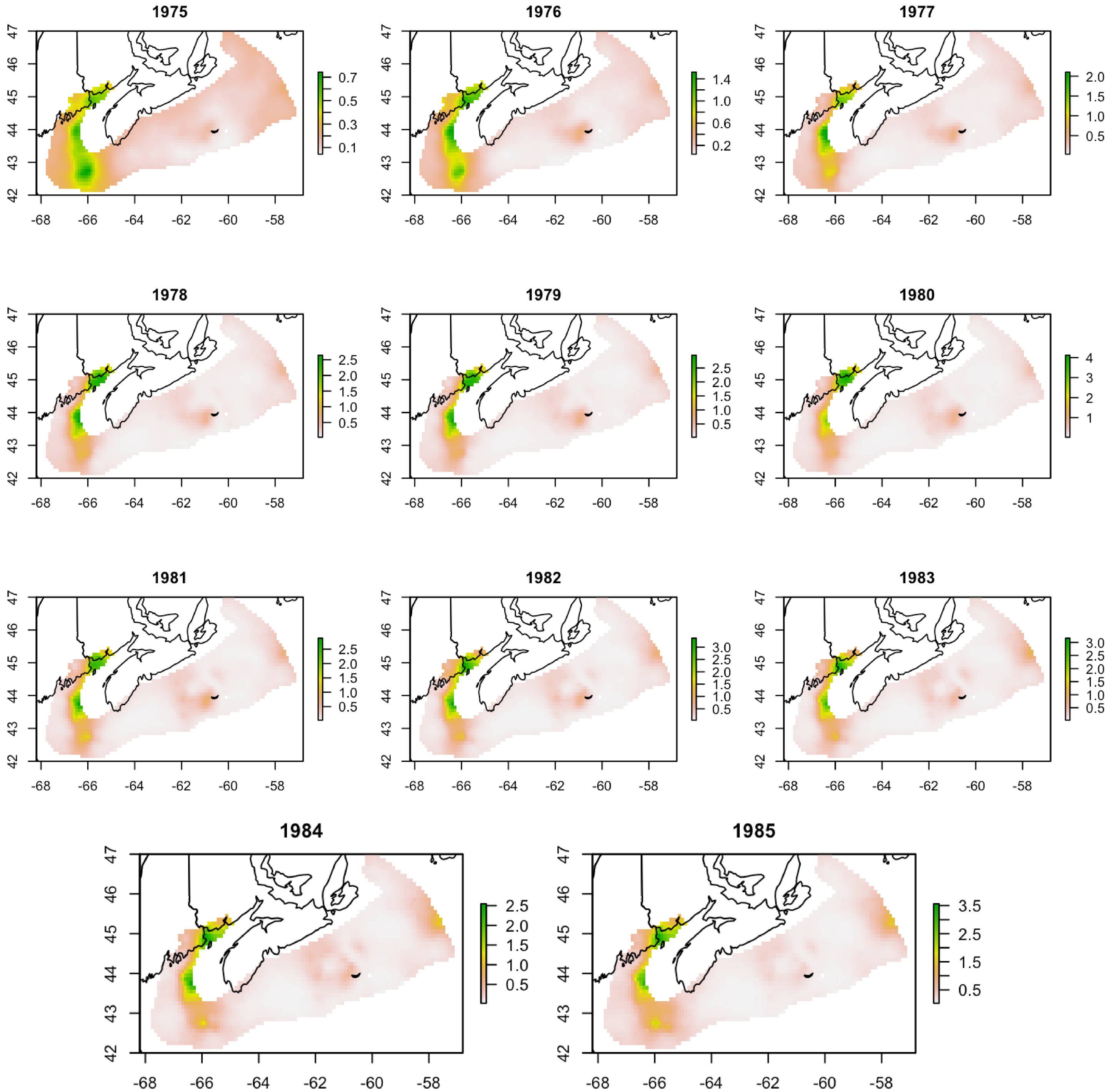

Fig. S3. Standard error of predicted historical smooth skate density. Shown is the calculated standard error for smooth skate density, 1975–1985

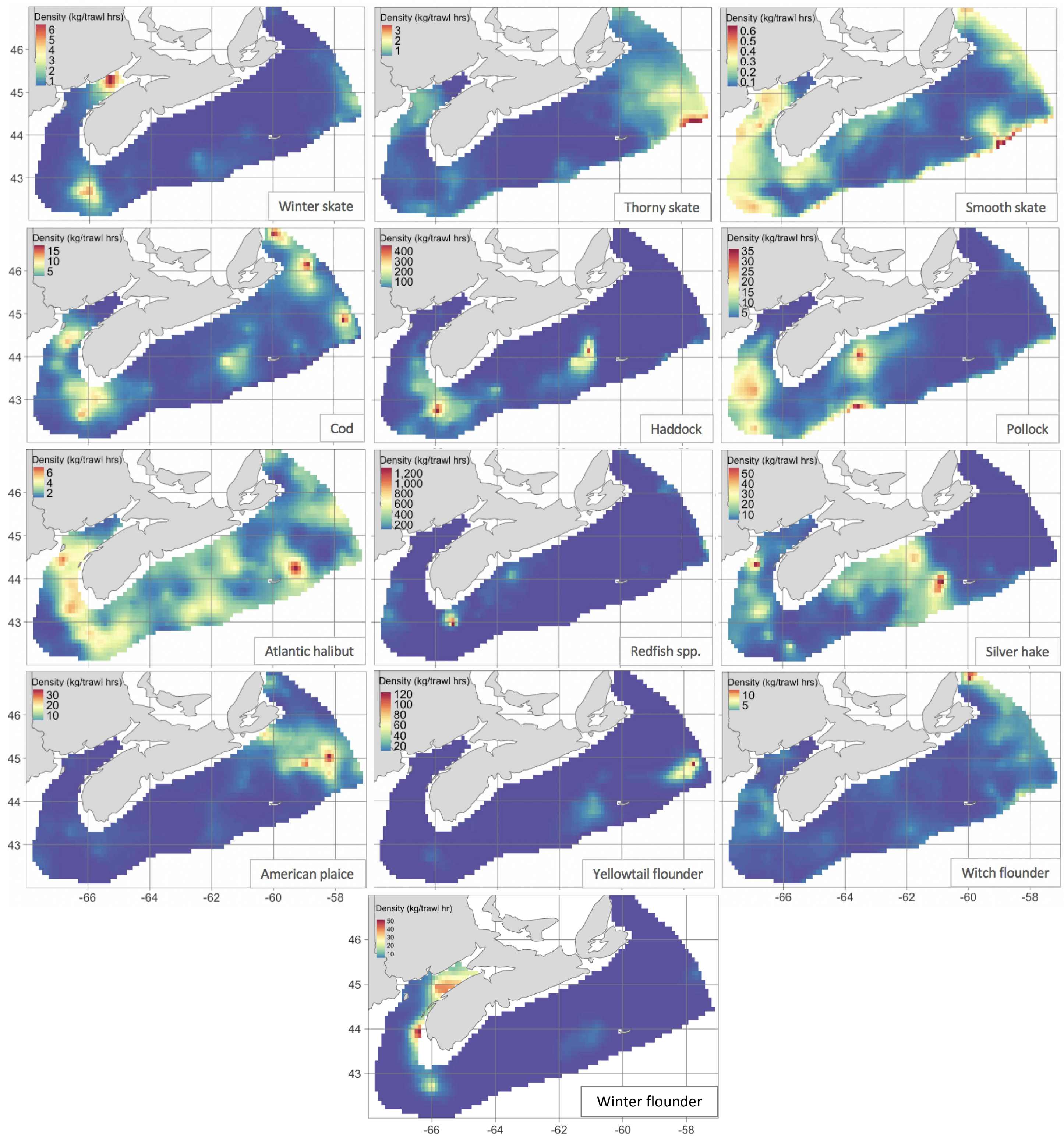

Fig. S4. Modelled mean species distribution. Relative density for each species predicted from RV survey data. Shown is mean species density for 2005–2015

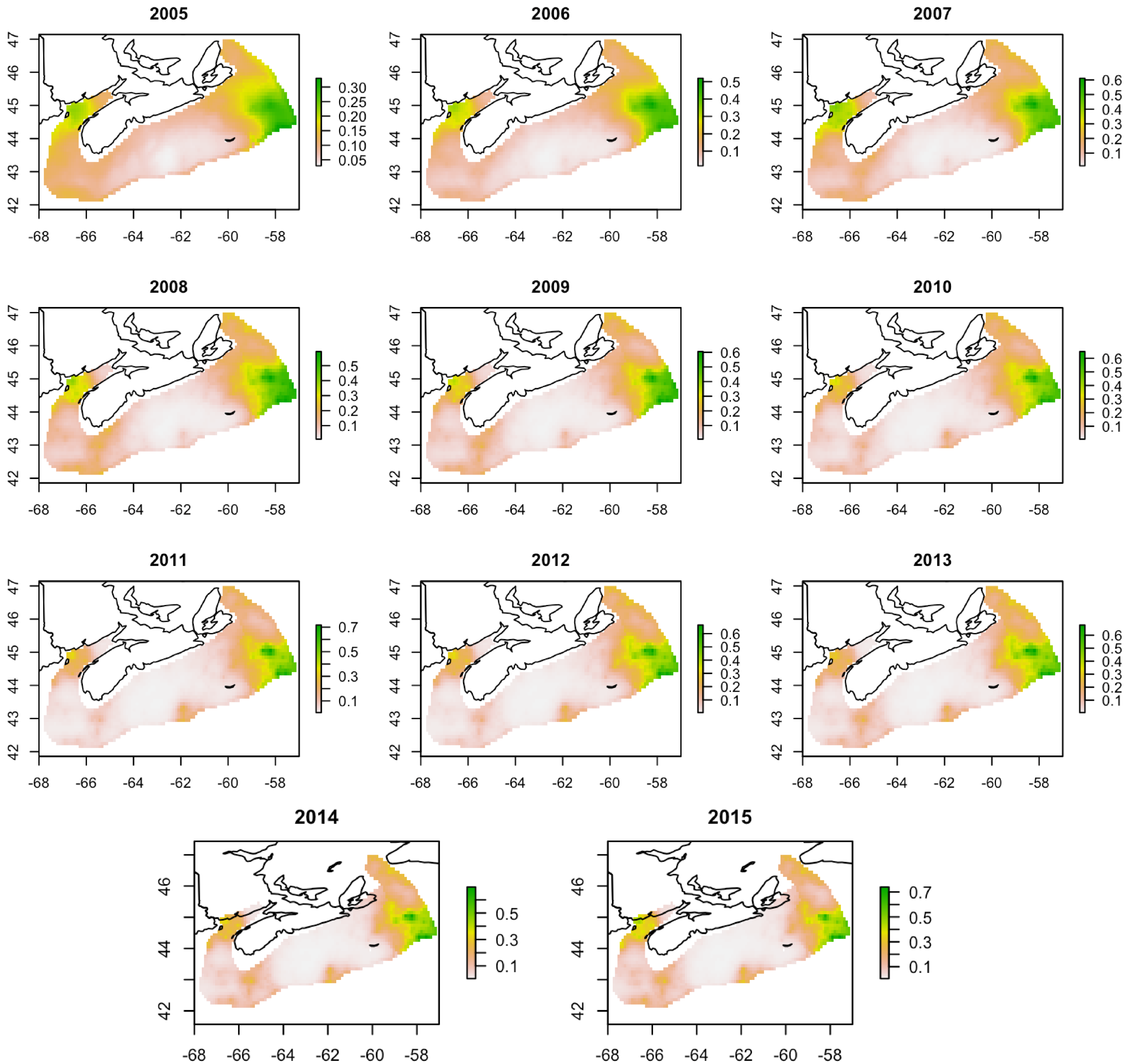

Fig. S5. Standard error of predicted thorny skate density. Shown is the calculated standard error for thorny skate, 2005–2015

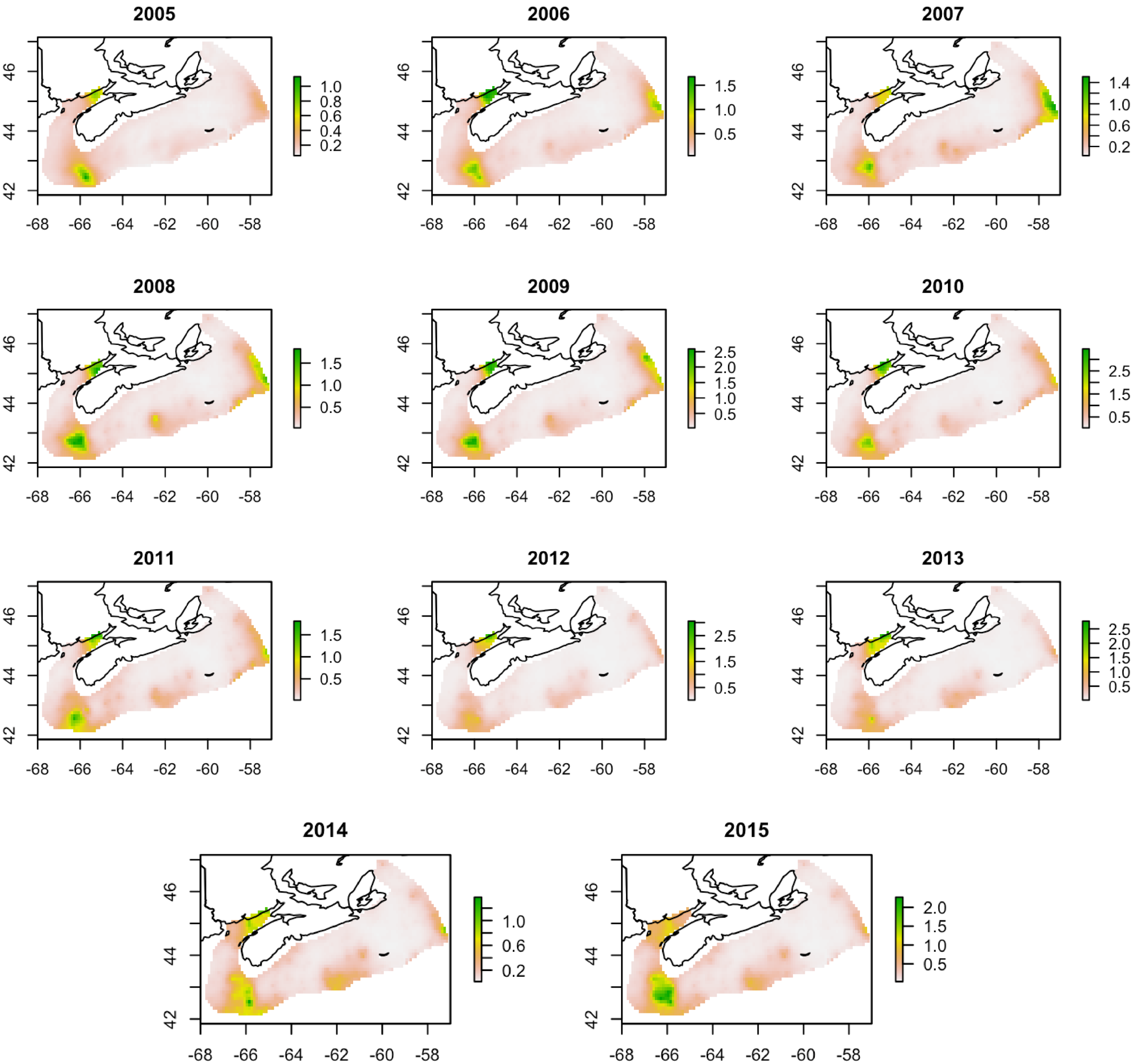

Fig. S6. Standard error of predicted winter skate density. Shown is the calculated standard error for winter skate density, 2005–2015

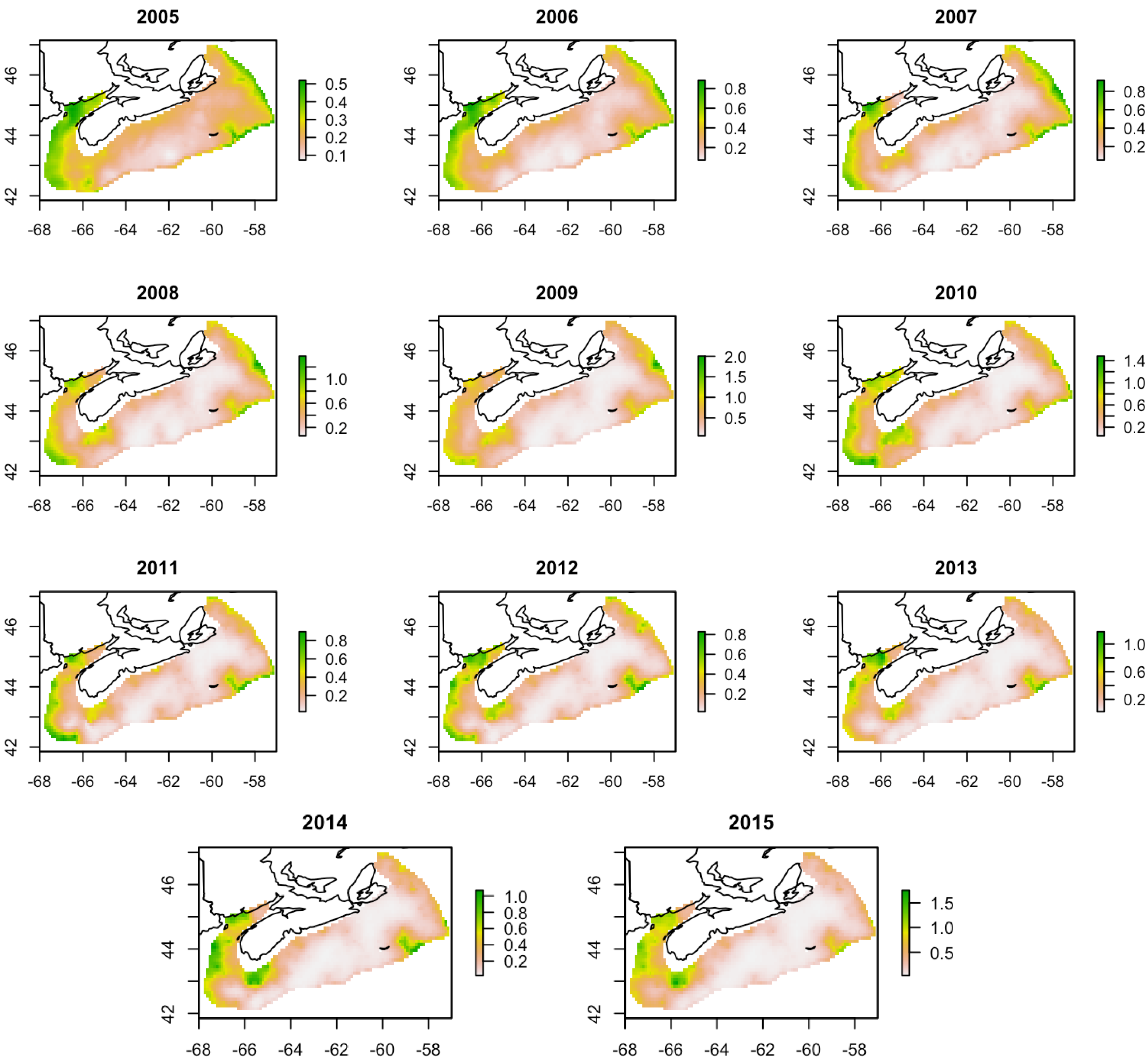

Fig. S7. Standard error of predicted smooth skate density. Shown is the calculated standard error for smooth skate density, 2005–2015

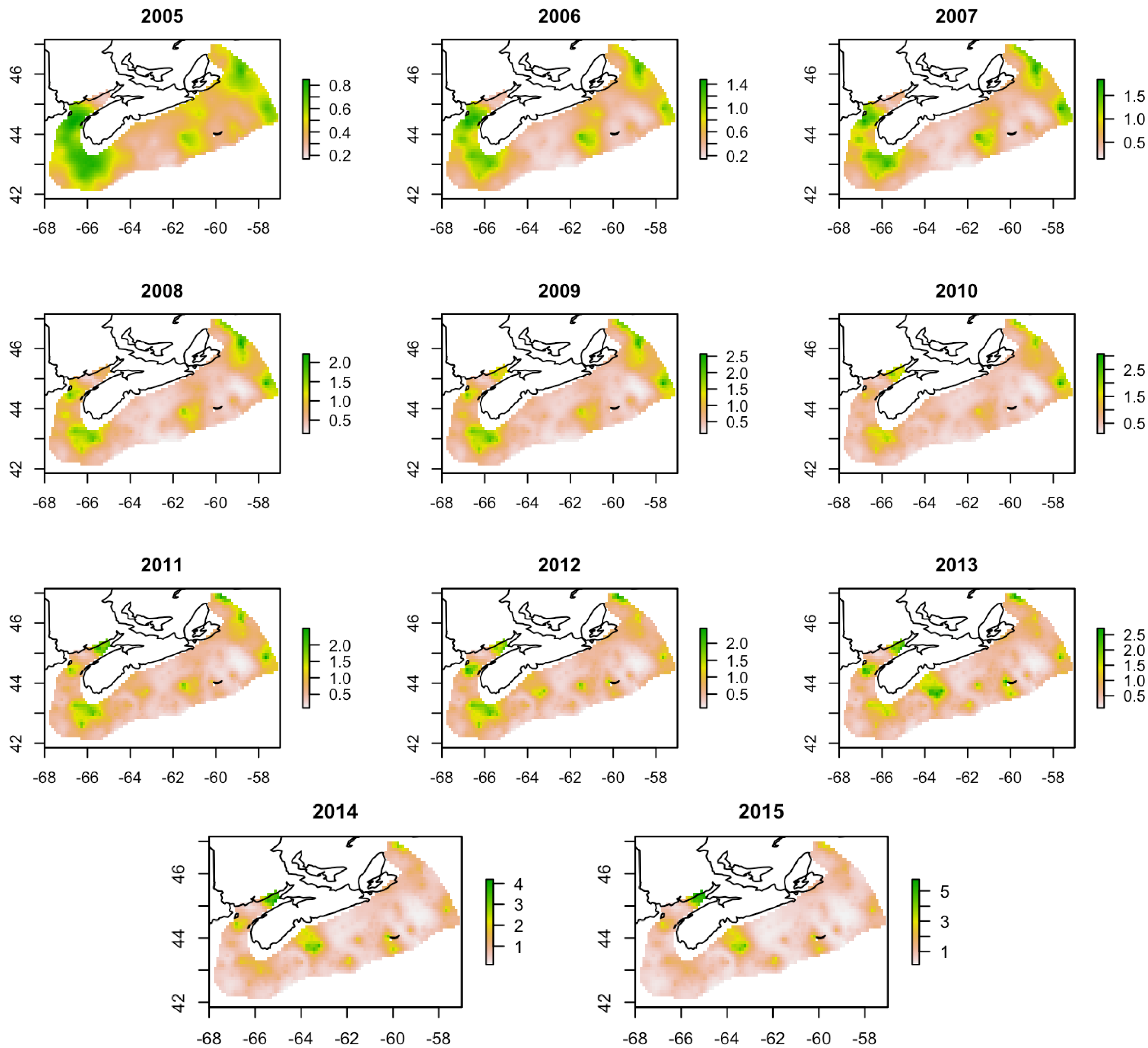

Fig. S8. Standard error of predicted cod density. Shown is the calculated error for cod density, 2005–2015

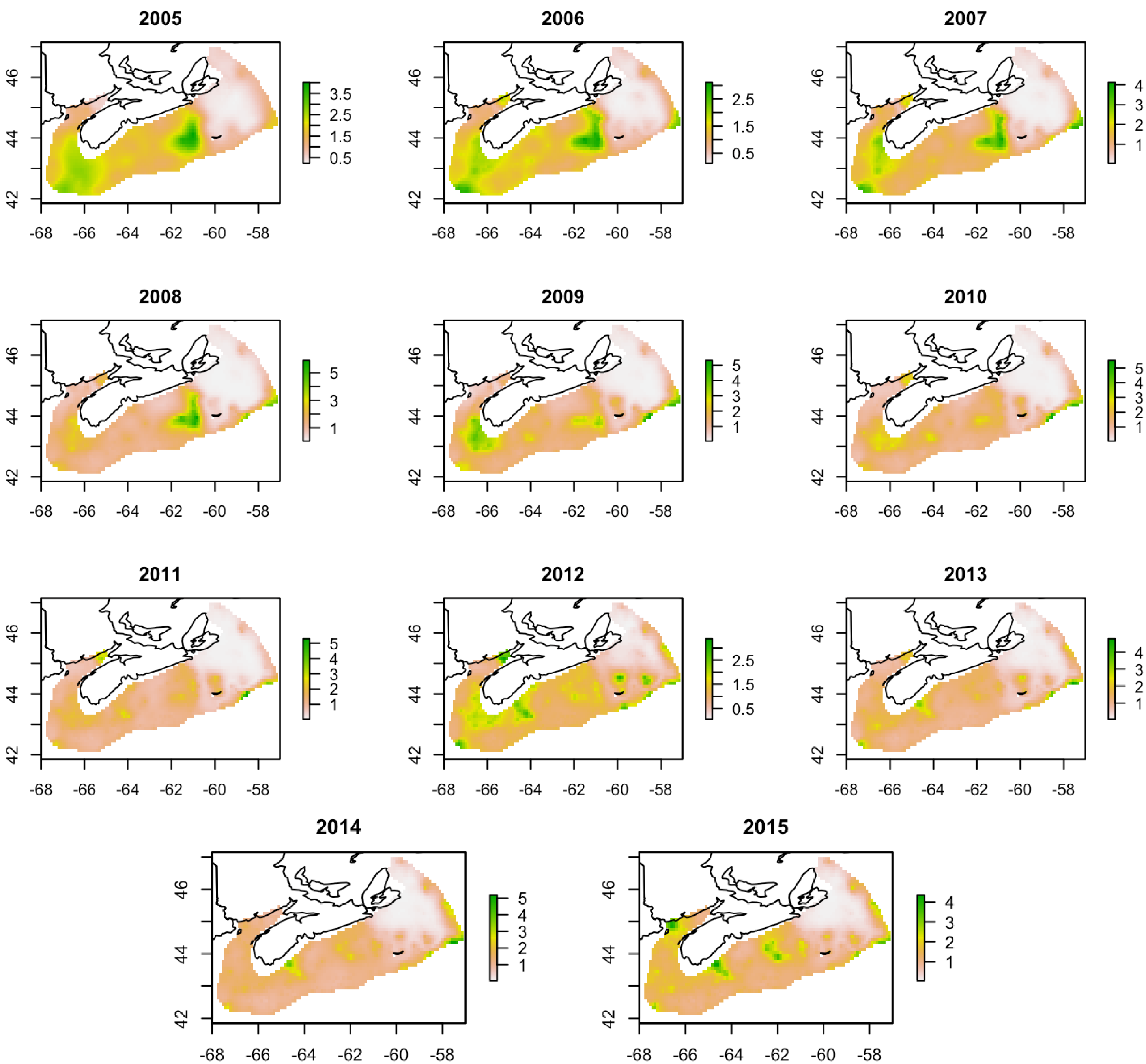

Fig. S9. Standard error of predicted haddock density. Shown is the calculated standard error for haddock density, 2005–2015

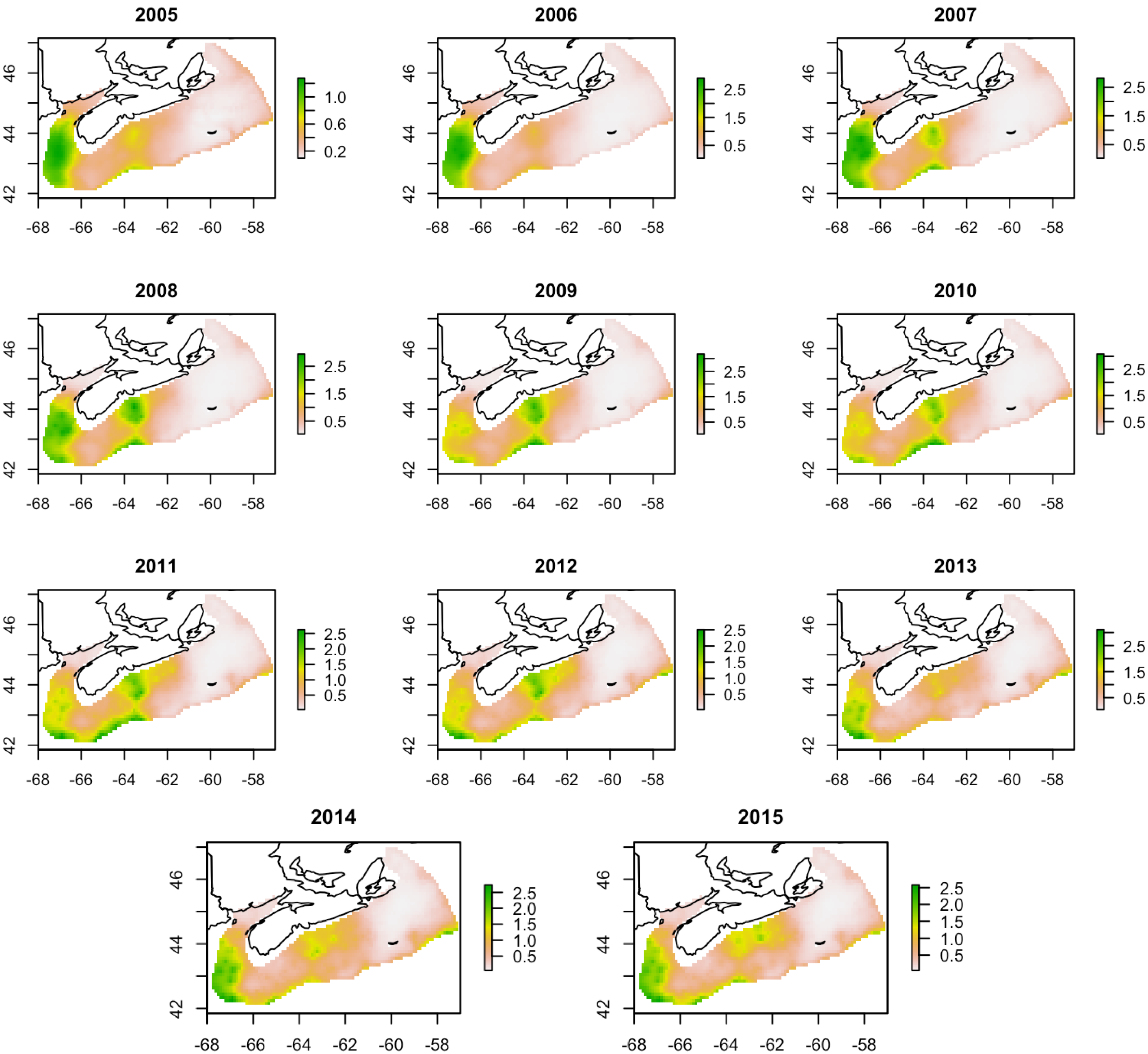

Fig. S10. Standard error of predicted pollock density. Shown is the calculated standard error for pollock density, 2005–2015

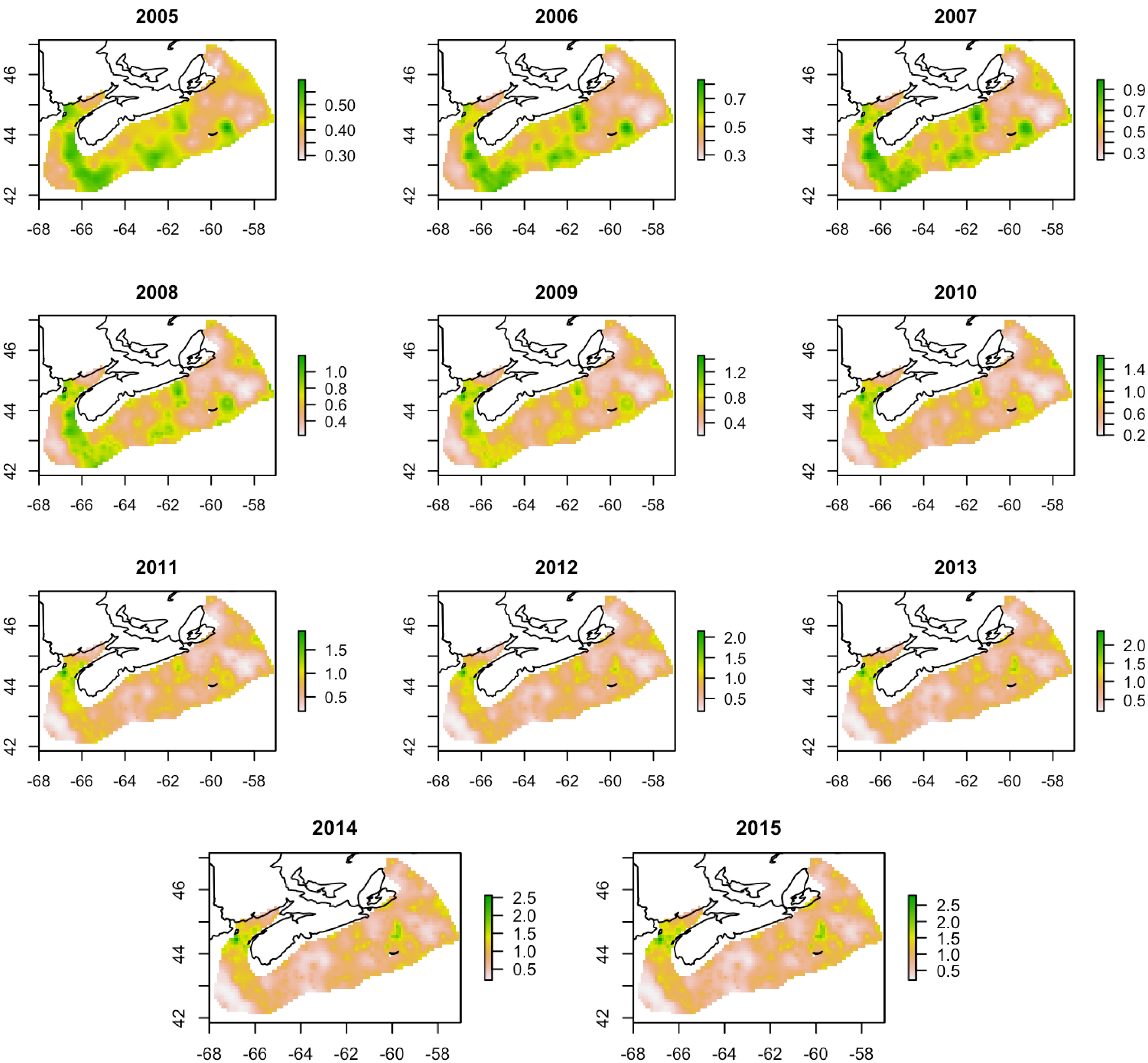

Fig. S11. Standard error of predicted Atlantic halibut density. Shown is the calculated standard error for Atlantic halibut density, 2005–2015

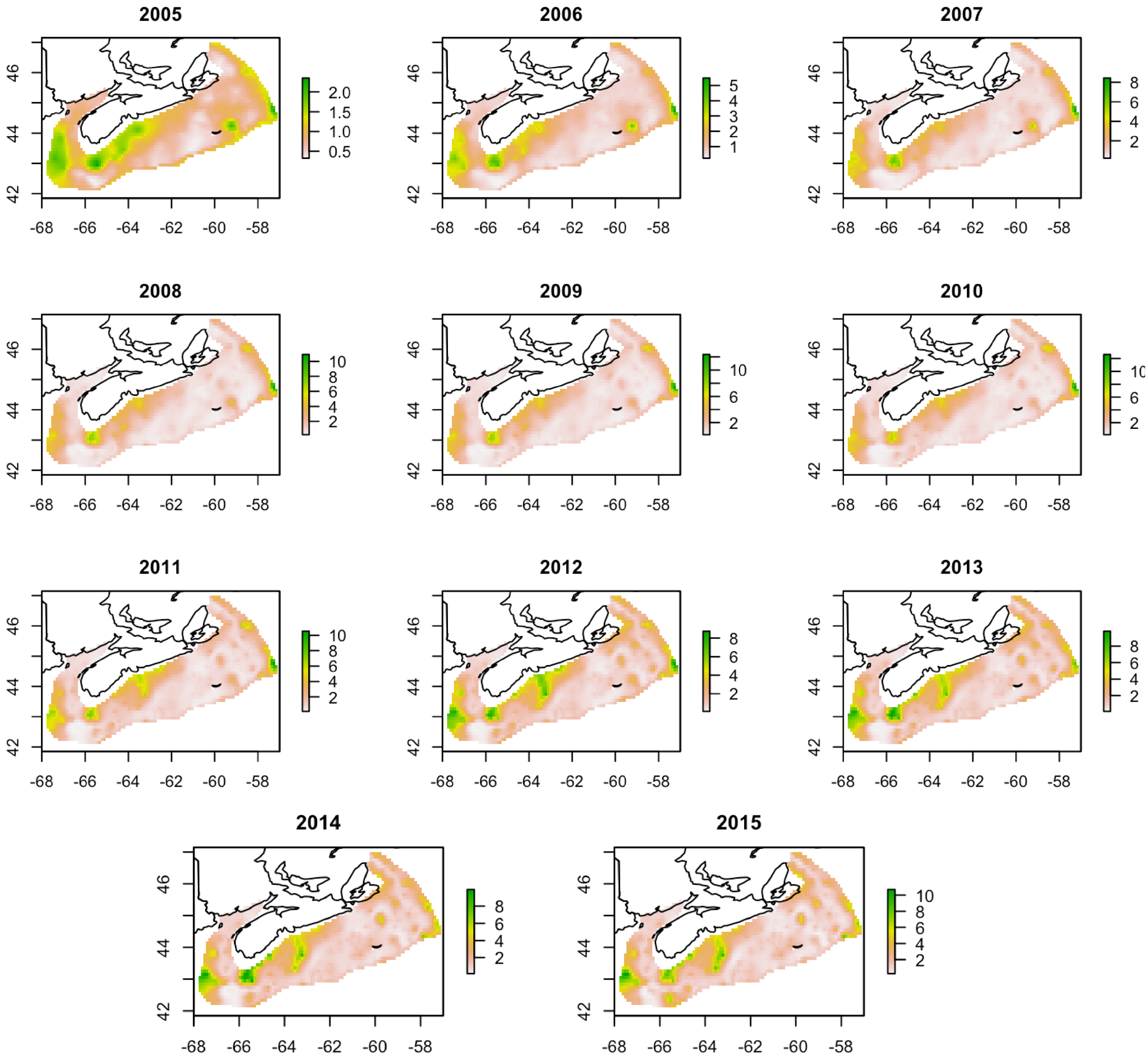

Fig. S12. Standard error of predicted redfish spp. density. Shown is the calculated standard error for redfish density, 2005–2015

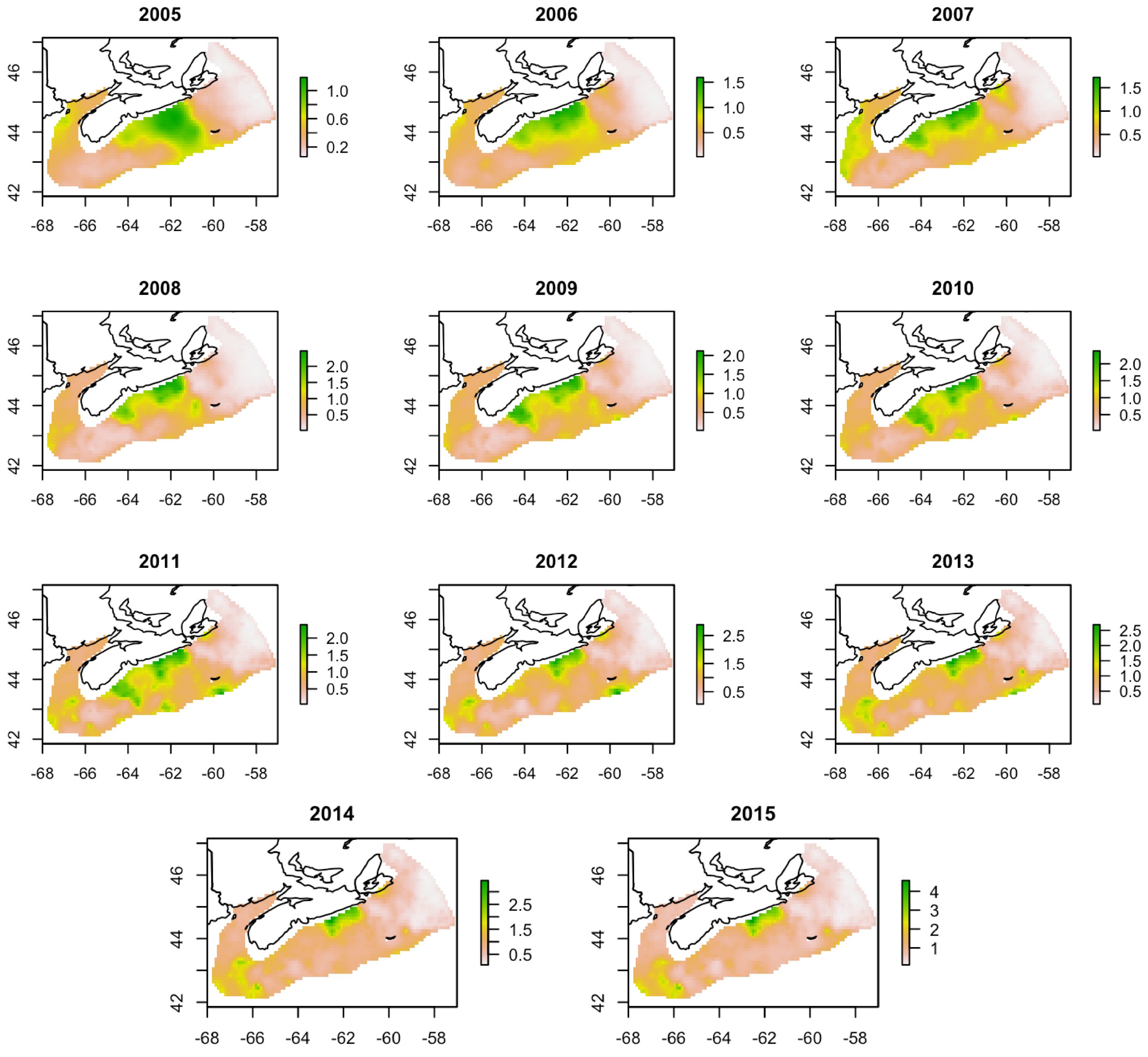

Fig. S13. Standard error of predicted silver hake density. Shown is the calculated standard error for silver hake density, 2005–2015

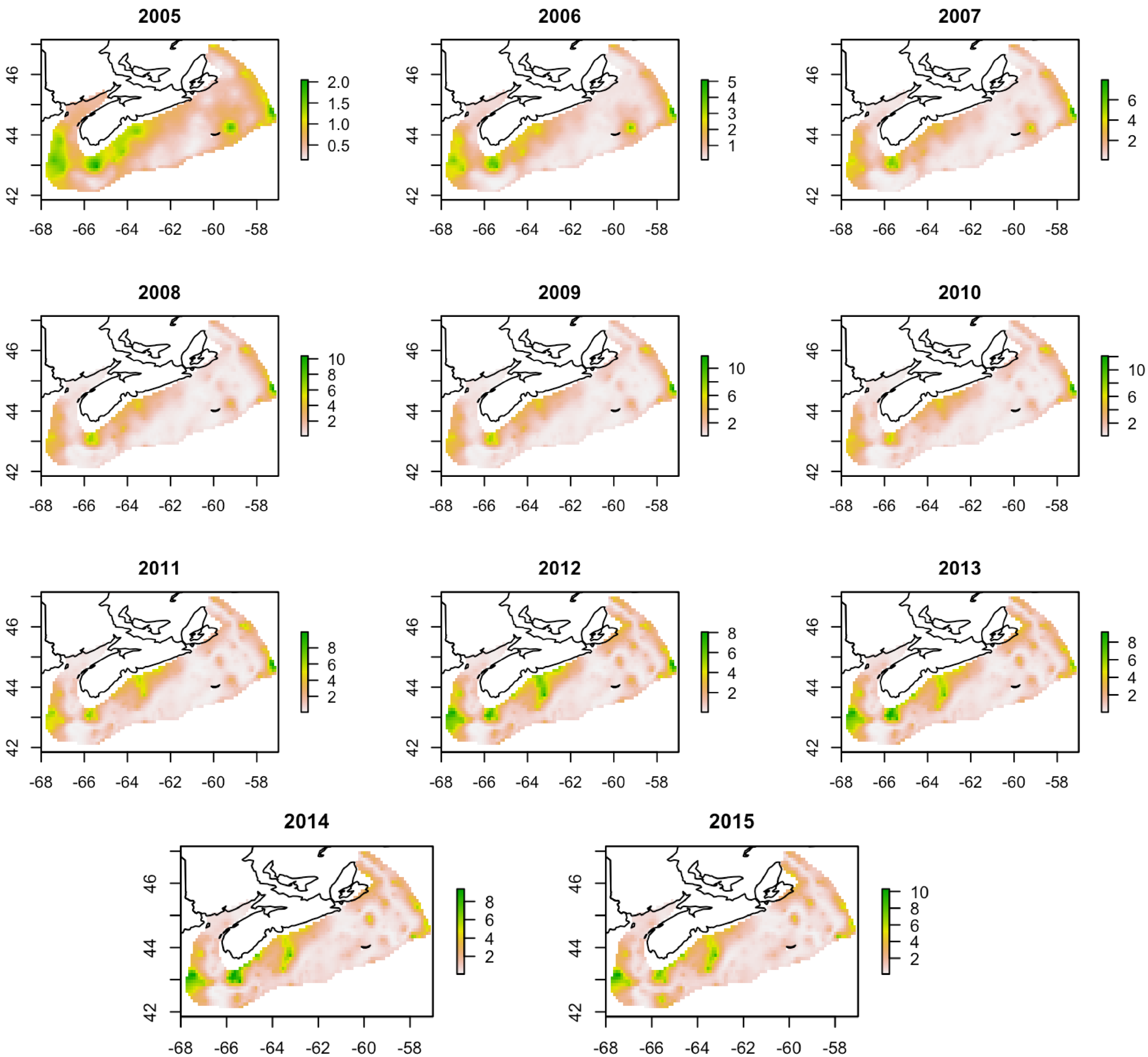

Fig. S14. Standard error of predicted American plaice density. Shown is the calculated standard error for American plaice density, 2005–2015

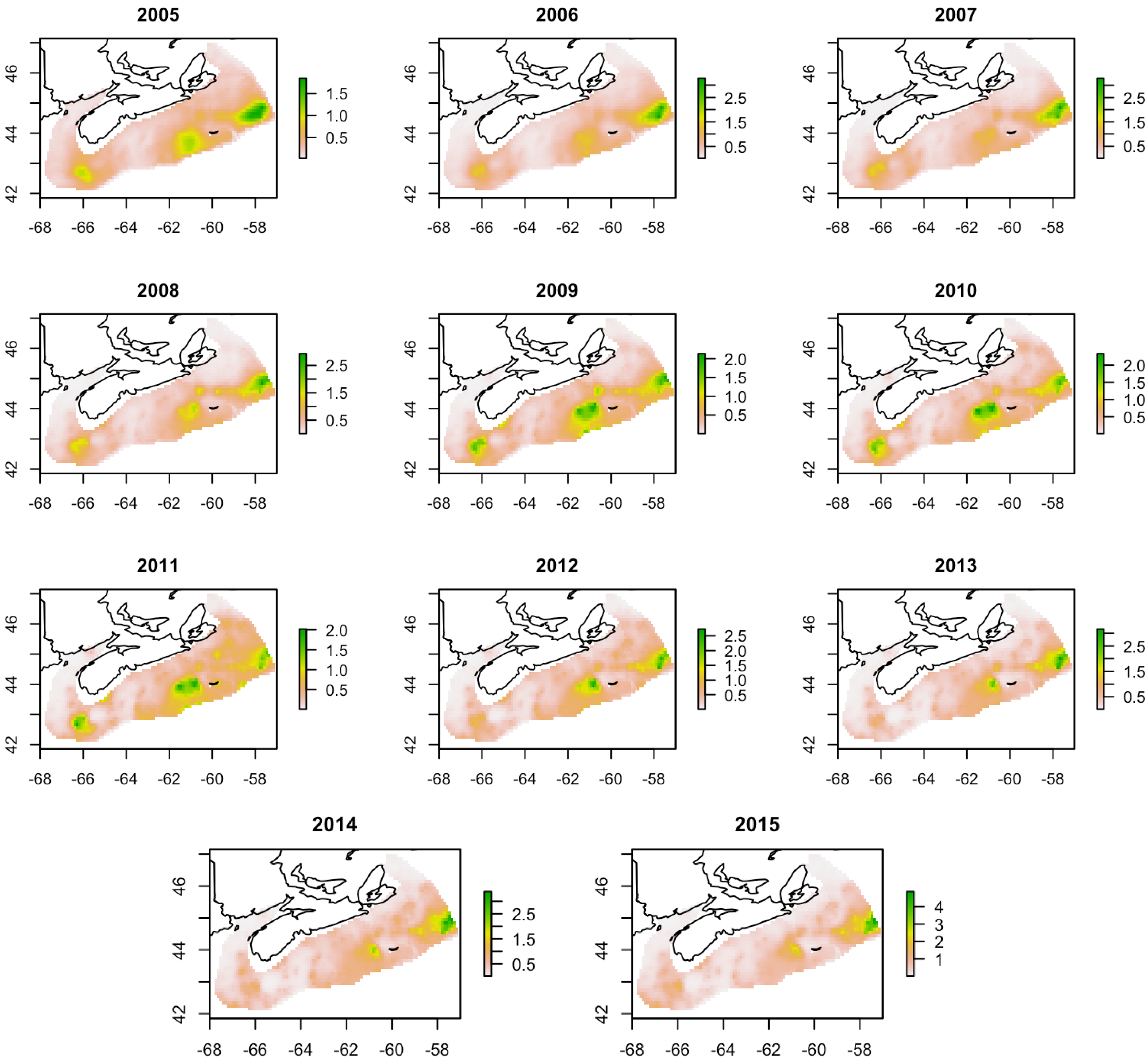

Fig. S15. Standard error of predicted yellowtail flounder density. Shown is the calculated standard error for yellowtail flounder density, 2005–2015

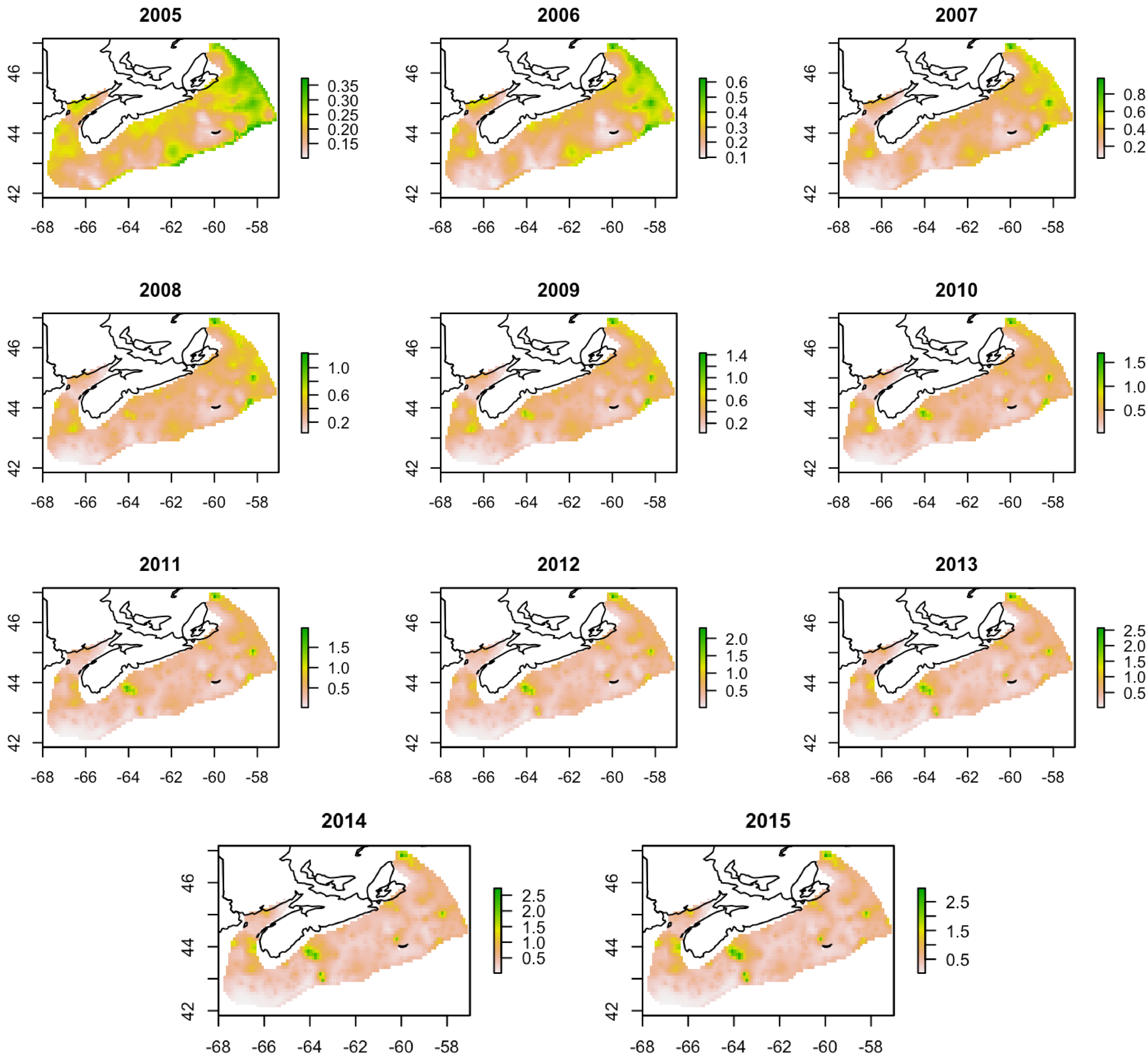

Fig. S16. Standard error of predicted witch flounder density. Shown is the calculated standard error for witch flounder density, 2005–2015

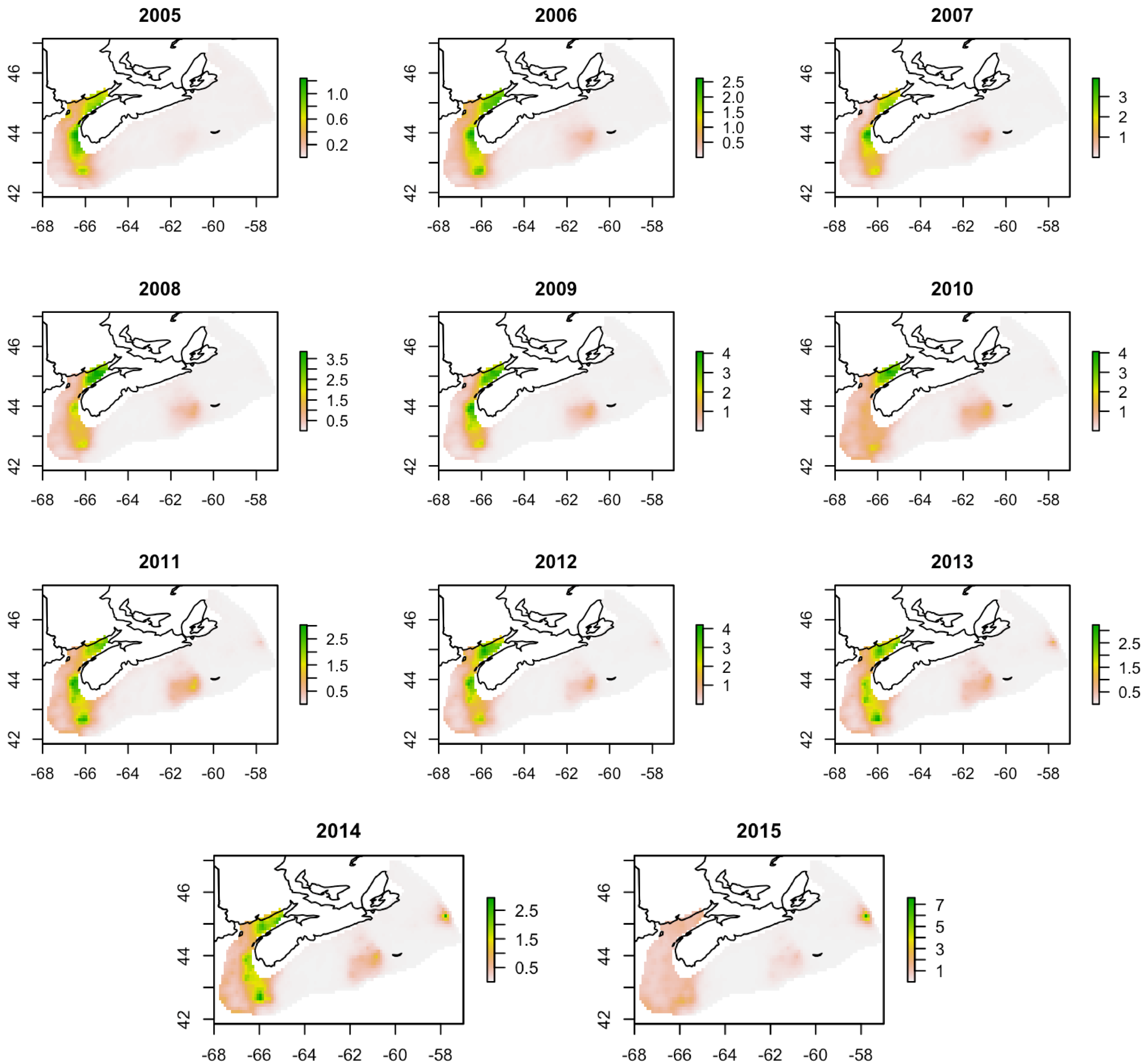

Fig. S17. Standard error of predicted winter flounder density. Shown is the calculated standard error for winter flounder density, 2005–2015

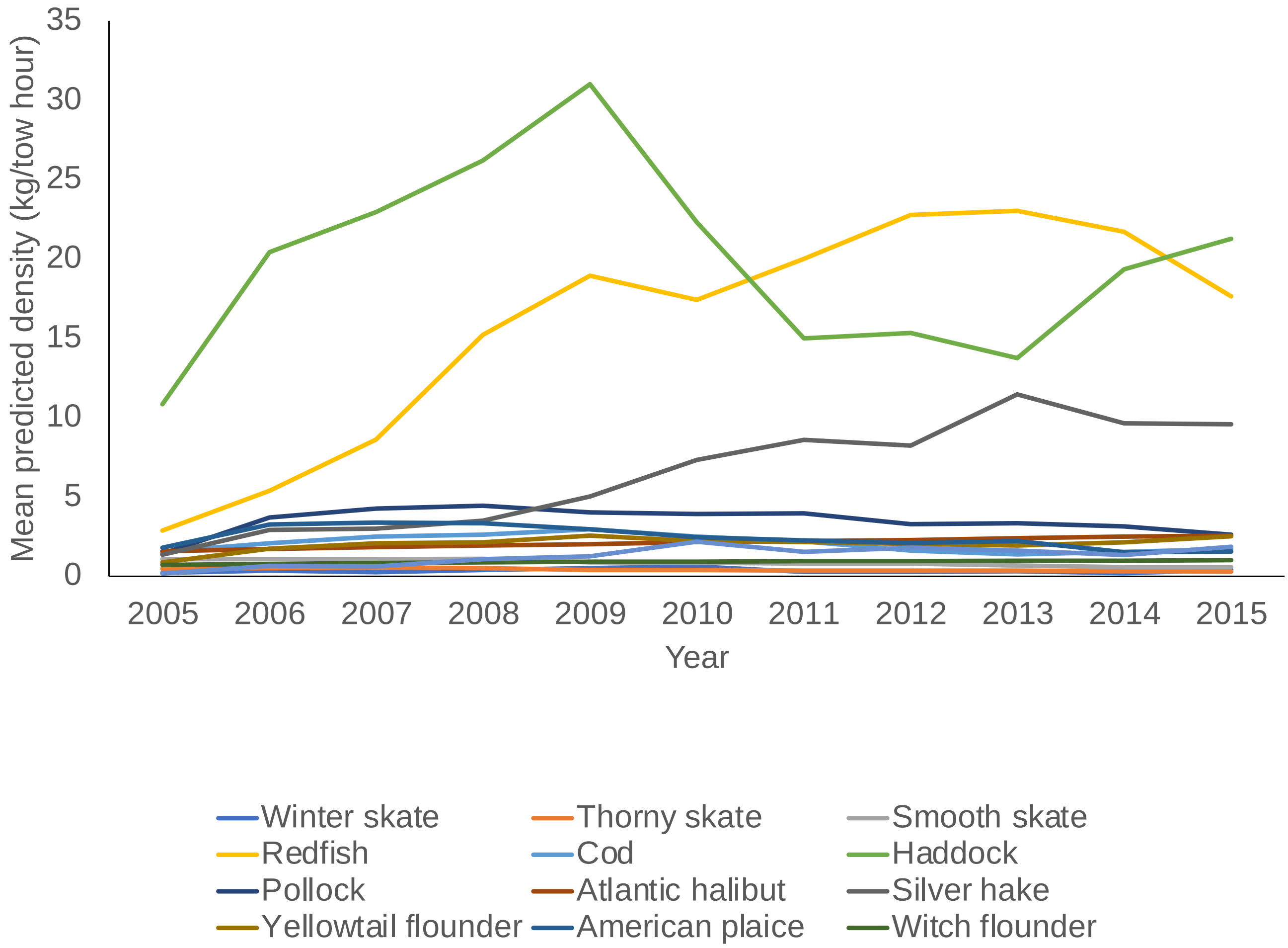

Fig. S18. Yearly means of species density responses. Curves show annual mean species presence multiplied by the exponent of predicted logCPUE (catch-per-unit-effort)