Supplementary Material for Hyman, Chiu, Fabrizio, and Lipcius

SUPPLEMENTARY DATA

Supplementary Figures

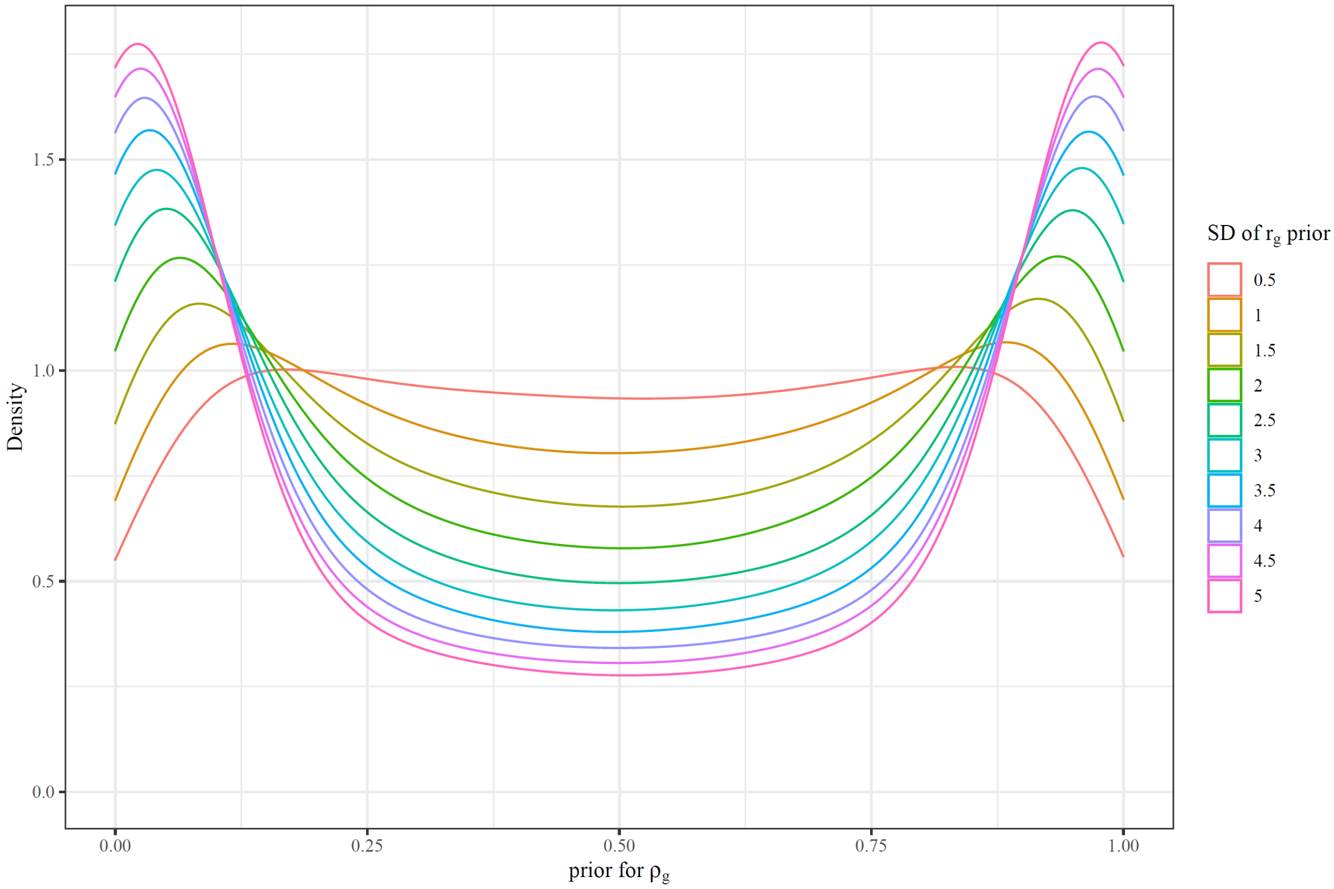

Figure S1. Marginal prior distributions of ρ_g with increasing standard deviations of the normally distributed prior for r_g (whose mean is 0), and a prior distribution of $U(0, 1)$ for P. The marginal prior distribution for ρ_g is approximately $U(0, 1)$ when a $N(0, 0.25)$ is imposed on r_g. Thus, constraining the prior for r_g to a relatively narrow distribution results in a diffuse marginal prior for ρ_g.

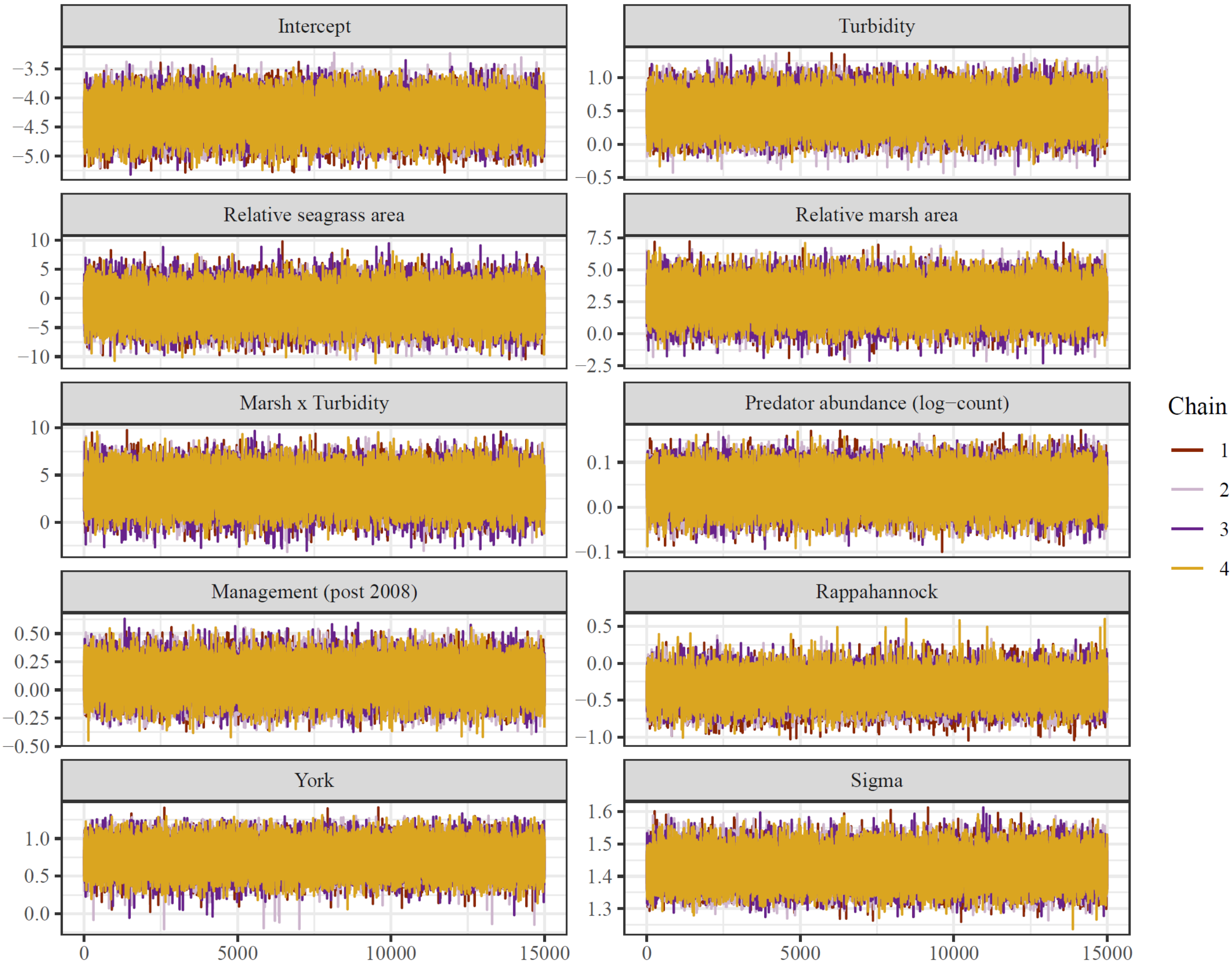

Figure S2. A set of trace plots for Model 4 parameters illustrating sampled values of each regression coefficient and σ_Φ per chain throughout the post burn-in iterations. Visual inspection of trace plots is used to evaluate convergence and mixing of the chains.

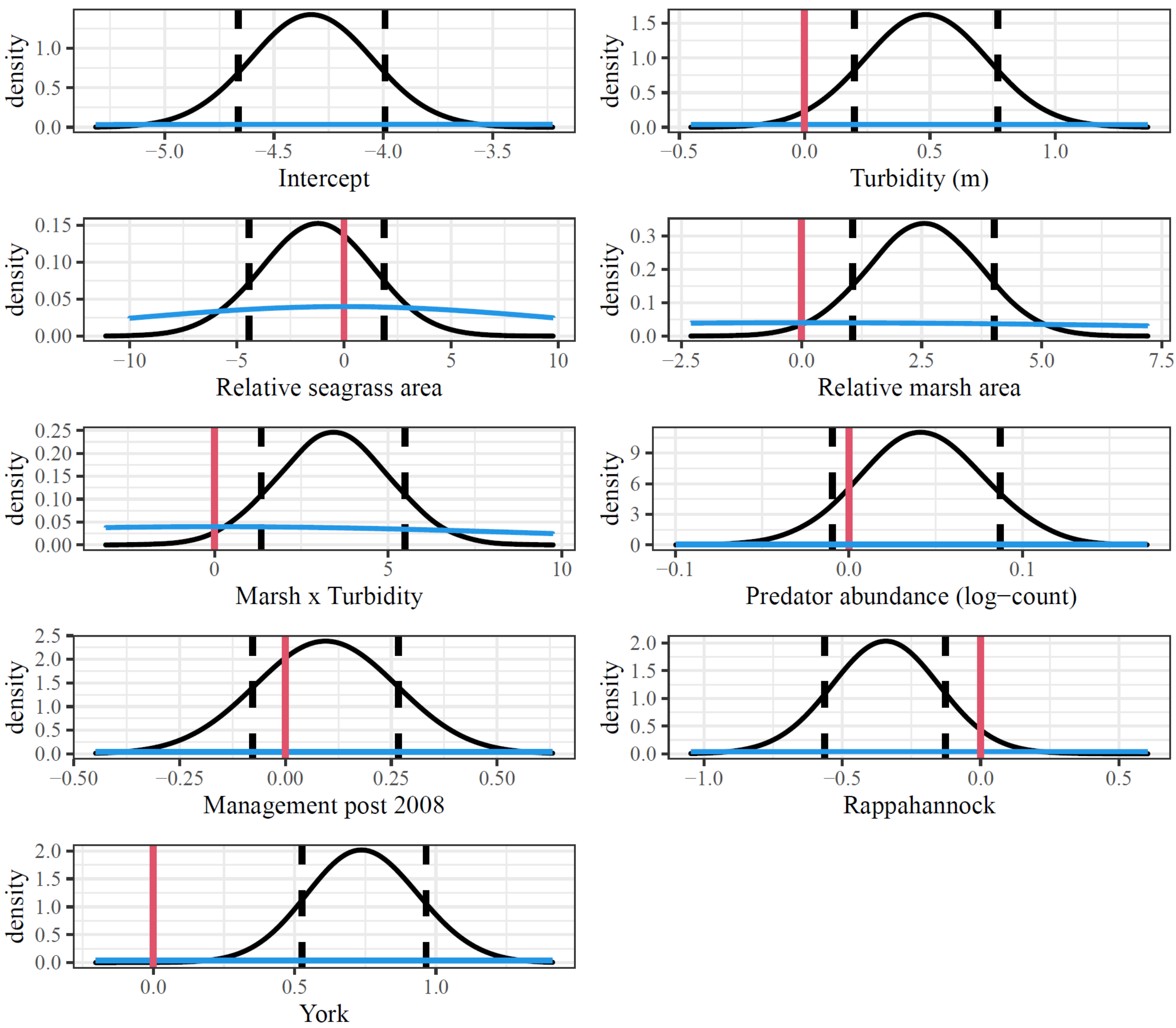

Figure S3. Posterior distributions (black) and prior distributions (blue) of regression coefficients from Model 4; dashed black lines denote 80% Bayesian confidence intervals, while solid red lines denote 0

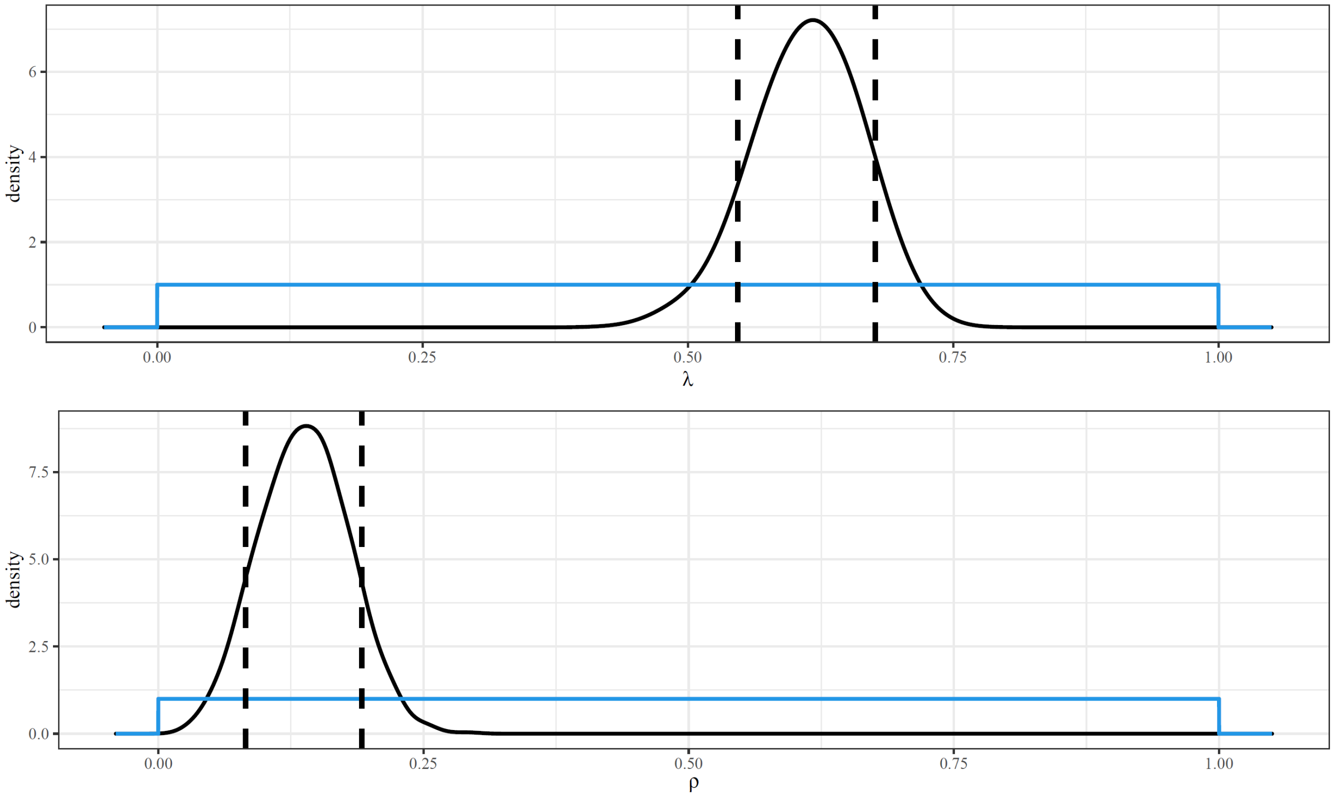

Figure S4. Posterior distributions (black) and prior distributions (blue) of autocorrelation parameters λ (spatial) and ρ (temporal) from Model 4; dashed black lines denote 80% Bayesian confidence intervals. Leave-future-out cross validation of Models 1–4 showed that the non-separable spatiotemporal dependence structure of Model 4 was necessary for good predictive performance, despite the small ρ.

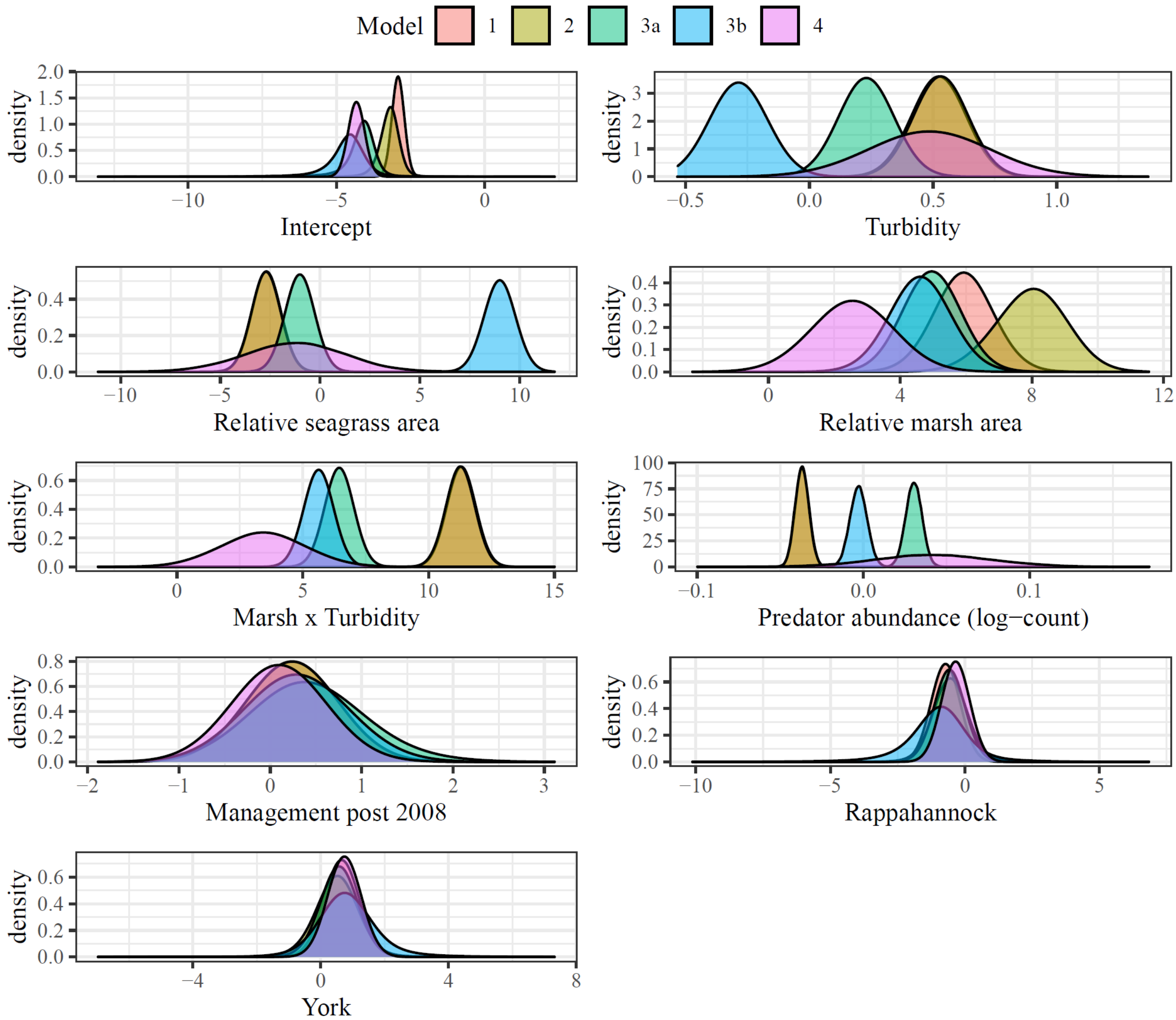

Figure S5. Posterior distributions of regression coefficients from Models 1-4.

APPENDICES

Appendix A: More on predictor variables

Mud: Unstructured habitat constitutes the majority of available shallow habitat in Chesapeake Bay, but varies considerably in food availability and predation refuge Lipcius et al. (2005). Evidence suggested unstructured mud may serve as an alternative nursery for juveniles where structurally complex habitat is unavailable due to relatively abundant alternative prey and potential for juveniles to bury deep in the soft substrate Mense and Wenner (1989); Rakocinski et al. (2003). Thus, in the earliest exploratory models, we had included mean percent mud composition of substrates in each section-year as a continuous covariate, in addition to those presented in Table 1. However, 80% Bayesian confidence intervals for the corresponding regression coefficient of this variable consistently included 0, and inclusion of the variable did not otherwise change inference results of the initial models. In contrast, the other variables, such as seagrass, management status, and predator abundance, were always kept in our models regardless of their statistical importance in explaining juvenile blue crab abundance, due to their implications on large-scale blue crab population management. Because percent mud composition did not carry the same implications with respect to management, it was excluded as a variable of interest in all models presented in this article.

Management: Although changes in the spatial distribution of fishing can impact abundances of economically important species (e.g., Carson et al., 2017), we employed a general before-after management term in lieu of a spatially-explicit term for the following reasons. First, although fishing pressure for adults in tributaries varies spatially, juveniles in the size fraction 20–40 mm CW, which were targeted in our study, are neither legal nor susceptible to the fishery due to their small size. For example, crab pot mesh size is large enough that small crabs (<40 mm CW) can easily escape. Moreover, since all mature female blue crabs in Chesapeake Bay migrate to the mouth of the estuary for spawning and larval release (Epifanio, 2019), spatial patterns in fishing pressure are decoupled from juvenile and postlarval abundance at sub-estuary scales.

Salinity: Juvenile blue crabs are tolerant of a wide range of salinity regimes ranging from nearly freshwater to hypersaline (Guerin and Stickle, 1992). Hence, it is unlikely that salinity alone would describe patterns in juvenile blue crab abundance. Rather, any changes in blue crab abundance along a salinity gradient are most likely reflective of changes in predator communities (e. g., Posey et al., 2005). Moreover, patterns in juvenile abundance along a turbidity gradient may be indicative of changes in predator efficiency. Due to the substantially collinear nature of salinity and turbidity, we chose to explicitly model turbidity and predator abundance and exclude salinity, which could better account for the biological patterns if present.

Appendix B: Defining areal units

Note that despite the arbitrary nature of areal unit definitions in practice, the one we employed in our work here did not meaningfully influence our results, or bias, our inference. In fact, initially, we explored numerous areal unit configurations when aggregating spatially random trawls. Alternative configurations included dividing each tributary into i) ten sections whose lengths were tributary dependent, ii) sections based on morphologically meaningful characteristics (e.g., branching structures and choke points), and iii) sections $\sim$2km in length along the tributary axis. In all cases, parameter estimates from models were practically identical. The final areal unit configuration was chosen based on the high number of areal units per year produced, and only a single section-year had 0 trawl tows.

Appendix C: Model inference, validation, and predictive performance

We reported statistical inference at the 80% Bayesian confidence level throughout this paper. The International Panel on Climate Change (IPCC), which employs Bayesian methods, designates $\geq$66% as "likely" and $\geq$90% as "very likely" (Chen et al., 2021). For practical implications, a confidence level between these values, e.g., 80%, therefore suggests a reliable range of parameter values.

Cross validation (CV) is a robust, generic method to adjudicate between competing statistical models. Unlike information theoretic criteria (e.g., AIC, BIC, DIC), cross validation assesses predictive performance directly by separating the data in a part that is used for fitting (i.e., training set) and another used to assess predictive adequacy (i.e., test set). Cross validation preference goes to the model that best predicts the out-of-sample test set withheld.

Cross validation is helpful in determining relative model generalizability. In a Bayesian CV framework, prediction intervals are computed using the posterior predictive distributions of the excluded values in the test set based on posterior distributions of model parameters to simulate the training set. Generalizability is determined based on the observed coverage, i.e., the proportion of excluded values which are successfully captured by their respective prediction intervals.

We used 80% prediction intervals to infer model performance of our suite of candidate models. Models yielding an observed coverage differing greatly from the nominal Bayesian predictive confidence level of 80% may indicate underfitting/overfitting. Cross validation results from Models 1, 2, 3a, and 3b, all indicated underfitting (being less complex than Model 4) and poor predictive performance. In contrast, posterior prediction intervals of Model 4 contained 81% of excluded data ($n = 36$), indicating overall superior predictive performance relative to all other candidate models. Hence, we selected Model 4 as the model which best represents our observed data as well as the most generalizable model.

In contrast with simpler models, Model 4 is characterized by greater uncertainty in posterior distributions of predictor coefficients as well as posterior predictive confidence intervals used in cross validation (Figs. 2 and S5). This is a frequent characteristic of models with increasing complexity. Complex models (with a larger number of unknown model parameters) lead to more uncertainty in the inference, whereas simpler models which are inadequate in capturing latent dependence processes would give incorrect inference, irrespective of the amount of uncertainty.

REFERENCES

Carson, S., Shackell, N., and Mills Flemming, J. (2017). Local overfishing may be avoided by examining parameters of a spatio-temporal model. *PloS one* 12, e0184427

Chen, D., Rojas, M., Samset, B., Cobb, K., Niang, A. D., Edwards, P., et al. (2021). Framing, context, and methods. In *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, eds. V. Masson-Delmotte, P. Zhai, A. Pirani, S. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. Matthews, T. Maycock, T. Waterfield, O. Yelekçi, R. Yu, , and B. Zhou (Cambridge: Cambridge University Press), chap. 1

Epifanio, C. E. (2019). Early life history of the blue crab *Callinectes sapidus*: a review. *Journal of Shellfish Research* 38, 1–22

Guerin, J. and Stickle, W. (1992). Effects of salinity gradients on the tolerance and bioenergetics of juvenile blue crabs (*Callinectes sapidus*) from waters of different environmental salinities. *Marine Biology* 114, 391–396

Lipcius, R. N., Seitz, R. D., Seebo, M. S., and Colón-Carrión, D. (2005). Density, abundance and survival of the blue crab in seagrass and unstructured salt marsh nurseries of Chesapeake Bay. *Journal of Experimental Marine Biology and Ecology* 319, 69–80

Mense, D. J. and Wenner, E. L. (1989). Distribution and abundance of early life history stages of the blue crab, *Callinectes sapidus*, in tidal marsh creeks near Charleston, South Carolina. *Estuaries* 12, 157–168

Posey, M. H., Alphin, T. D., Harwell, H., and Allen, B. (2005). Importance of low salinity areas for juvenile blue crabs, *Callinectes sapidus Rathbun*, in river-dominated estuaries of southeastern United States. *Journal of Experimental Marine Biology and Ecology* 319, 81–100

Rakocinski, C. F., Perry, H. M., Abney, M. A., and Larsen, K. M. (2003). Soft-sediment recruitment dynamics of early blue crab stages in Mississippi Sound. *Bulletin of Marine Science* 72, 393–408