Canadian Science Publishing
NRC Research Press journals | Publishing Services

Canadian Journal of Fisheries and Aquatic Sciences
Journal canadien des sciences halieutiques et aquatiques

Spatio-temporal index standardization improves the stock assessment of northern shrimp in the Gulf of Maine

Journal:	*Canadian Journal of Fisheries and Aquatic Sciences*
Manuscript ID	cjfas-2016-0137.R4
Manuscript Type:	Article
Date Submitted by the Author:	22-Feb-2017
Complete List of Authors:	Cao, Jie; University of Maine, school of marine sciences Thorson, James T.; Northwest Fisheries Science Center, National Marine Fisheries Service, NOAA Richards, R. Anne; Northeast Fisheries Science Center, National Marine Fisheries Service Chen, Yong; University of Maine,
Keyword:	abundance index, delta-generalized linear mixed model, Gaussian random field, spatial modeling, size-structured assessment model

Spatio-temporal index standardization improves the stock assessment of northern shrimp in the Gulf of Maine

Jie Cao[1*]

James T. Thorson[2]

R. Anne Richards[3]

Yong Chen[1]

[1]School of Marine Sciences, University of Maine, Orono, Maine 04469, USA

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

[3]Northeast Fisheries Science Center, 166 Water St., Woods Hole, MA 02543

Correspondence Email: jie.cao@maine.edu

ABSTRACT

Estimated trends in relative stock abundance are a primary input to fish stock assessments. Accurate and precise estimates are essential for successful conservation and management. Scientifically designed data collection ensures that estimates of relative abundance are unbiased. However, the statistical efficiency of a design-based estimator may be low under certain circumstances. We apply a recently developed spatio-temporal model that incorporates habitat variables to estimate a model-based abundance index for northern shrimp (*Pandalus borealis*) in the Gulf of Maine. We contrast this spatio-temporal index with a classical design-based index and evaluate the impacts of differences between the two abundance indices on the stock assessment. We show that using the spatio-temporal index in the assessment model greatly alters the estimates of recruitment and spawning stock biomass and the determination of stock status. Also, incorporating the spatio-temporal index leads to less retrospective bias and outperforms the model with design-based index in terms of predictive performance through a retrospective cross-validation test. Our results suggest that temporal variability of population abundance could be exaggerated by the design-based estimator and such imprecision may greatly affect the performance of a stock assessment and subsequent development of management decisions.

Keywords: abundance index, delta-generalized linear mixed model, Gaussian random field, spatial modeling, size-structured assessment model

INTRODUCTION

Periodic fishery independent surveys provide important information for fisheries stock assessment and management. Abundance indices are often the primary information derived from the surveys, and are essential to perform an adequate assessment. Contemporary assessments largely rely upon abundance indices which are assumed to be proportional to population abundance (Maunder and Punt 2004). The precision of these estimates is critical in influencing uncertainty associated with stock assessment and subsequent development of management decisions (e.g., total allowable catches).

There are mainly two types of methods used to estimate abundance indices from fishery independent surveys: classical design-based estimators and model-based estimators. Fishery independent surveys are generally well designed statistically with randomized sampling locations. They usually have stratified random design as appropriate stratification can increase the precision of estimates with limited sampling effort (Cao et al. 2015). The design-based estimators infer the population abundance according to the randomness induced by the sampling design. For example, the commonly used stratified-random design estimator generates the abundance index as the stratified mean for each stratum weighted by area (Smith 1990). However, model-based approaches analyze the survey data conditional on a hypothesized statistical model to control for confounding effects (e.g., differences in survey catchability; Helser et al. 2004; Thorson and Ward 2014) and make inference according to an assumed probability function for the response variable (Chen et al. 2004). These two different estimators have different philosophies of statistical inference (Smith 1990). For design-based theory, inferences about population quantities are based on assuming they are fixed, whereas application of models assumes that there is an underlying stochastic process generating the data (Smith

1990). Both design- and model-based estimators are commonly used in the United States (Helser et al. 2004; Thorson et al. 2015).

Conventional models could produce biased estimates of abundance (Ye and Dennis 2009). However, spatio-temporal models developed recently have been demonstrated to produce more precise and accurate abundance indices than either design-based or conventional model-based approaches (Shelton et al. 2014; Thorson et al. 2015). A spatio-temporal model can account for spatial dependence which results in estimating a smoothed surface representing spatial variation in density (Thorson et al. 2015). Also, habitat variables, e.g. (depth, bottom substrate type, temperature and salinity), can be incorporated into the spatio-temporal model as covariates. This can potentially lead to more precise estimates of abundance, especially when the underlying population distribution is largely dependent on habitat variables. In contrast, conventional design-based approaches cannot explicitly incorporate habitat variables and may produce imprecise estimates of abundance, particularly in the situation where habitat preference occurs and when strata included in the sampling design do not capture a large portion of spatial variation in density. Shelton et al. (2014) showed that the habitat preference of darkblotched rockfish in selected sampling locations can largely explain the variation in survey catch rates. Therefore, the temporal variability of population abundance may be exaggerated by a design-based estimator when the randomized sampling locations happen to fall in good habitat for some years, and vice versa (Shelton et al. 2014).

The spatio-temporal model has been applied to data for 28 groundfish species off the U.S. West Coast and the results were compared to a conventional model-based approach (Thorson et al. 2015). In general, the spatio-temporal and stratified indices showed similar trends (Thorson et al. 2015), while the uncertainty associated with the stratified index was substantially larger than

that associated with the spatio-temporal index (Thorson et al. 2015). In some cases, a statistically inefficient annual estimate from a design-based estimator (i.e., spikes in abundance index with high expected imprecision) could be avoided by using a spatio-temporal model (Shelton et al. 2014). However, few studies have shown the impacts of disparity between spatio-temporal and design-based indices on stock assessment results and performance. Given that the abundance index provides primary information for stock assessment, such studies are needed to better understand the practical importance of spatio-temporal index standardization.

We implement the spatio-temporal model to data collected from a summer shrimp bottom trawl survey designed specifically for monitoring northern shrimp (*Pandalus borealis*) in the western Gulf of Maine (GOM), and results are compared with an existing design-based index used in the stock assessment. Habitat variables are included in the spatio-temporal model to evaluate whether inclusion of habitat covariates helps explain the distribution of this species. We then estimate parameters for a stock assessment model for northern shrimp using the spatio-temporal index and compare the assessment model performance and outputs with those obtained based on the design-based index. Northern shrimp serves as our case study for two reasons: (1) there is high temporal variation observed in its design-based index, including an unlikely high spike in 2006; and (2) northern shrimp are considered to be sensitive to environmental changes (Richards et al. 2012) and there might be high inter-annual variation in their spatial distribution. This paper presents a real fishery example to demonstrate that spatio-temporal index standardization can improve the performance of a stock assessment model.

METHODS

Spatio-temporal delta generalized linear mixed model

105 A zero-inflated generalized linear mixed modeling framework was used in this study.
106 This framework explains the catch (in numbers) as function of two processes: (1) the probability
107 of sampling habitat where the species is present; and (2) the distribution of catches in habitat
108 where the species is present. Sampling in unoccupied habitat will always generate a catch of
109 zero (a "true zero"), while sampling in occupied habitat may also generate a catch of zero (a
110 "false zero"; see Martin et al. (2005)). Specifically, catches were assumed to follow a zero-
111 inflated negative binomial distribution:

112 $$\Pr[C = c|C > 0] = \begin{cases} (1-p) + p \times \mathrm{NB}(C|n,r) & \text{if } C = 0 \\ p \times \mathrm{NB}(C|n,r) & \text{if } C > 0 \end{cases} \quad (1)$$

113 where p is the probability of sampling in where the species is present (so $1-p$ is the probability
114 of a "true" zero), and $\mathrm{NB}(c|n,r)$ is the negative binomial probability density function evaluated
115 at value c with size n and probability r. We specified a quadratic function for the variance of the
116 negative binomial distribution as a function of the mean:

117 $$\sigma^2 = (1+\theta_1)\lambda + \theta_2\lambda^2 \quad (2a)$$

118 where λ is the expected catch in occupied habitat, and where this variance was then used to
119 calculate the size and probability parameters (r and p) for the negative binomial distribution:

120 $$r = \lambda/\sigma^2 \quad (2b)$$

121 $$n = \lambda r/(1-r) \quad (2c)$$

122 This distribution involves estimating the probability of sampling occupied habitat (p), the
123 expected catch in occupied habitat (λ), and two variance parameters (θ_1 and θ_2), and its
124 expectation is then easily calculated ($\mathbb{E}(C) = p \times \lambda$).

125 We accounted for spatial dependence in both the probability of occupied habitat (p) and
126 the density given occupied habitat (λ). Accounting for spatial dependence can lead to better
127 inference, superior prediction, and a more accurate characterization of the variability of estimates,

and we did so in this model using Gaussian Markov random fields. Random fields describe random processes defined over parameter spaces with multiple dimensions. For example, in fisheries science the spatial parameter might represent variation in population density over two dimensions (latitude and longitude). Specifically, spatial dependence can be imposed by modeling a zero-mean stationary Gaussian random field, which defines the expected value, variance, and covariance of a multivariate realization from a stochastic process. For a zero mean stationary Gaussian random field w, the value of w at a given location s = (x, y) (where x and y are the easting and northing for that location) follows a normal distribution and the value of w at a finite number of locations follows a multivariate normal distribution:

$$\mathrm{w}[s] \sim MN(0, \sum \Theta), \quad (3)$$

where *MN* is a multivariate normal distribution, and $\sum \Theta$ is the covariance matrix of the two dimensional normal density. We specified that the covariance follows a Matern function (with smoothness ν =1), which measures spatial proximity in terms of distances between the locations. The Matern function is slightly more smooth than the exponential correlation function used in other recent spatio-temporal models in fisheries science (e.g., Kristensen et al. 2014). While we assumed that the random field is stationary, we included the potential impact of geometric anisotropy in order to deal with the situation that dependence may be different in different directions:

$$\sum(s, s') = \sigma_E^2 \cdot Matern(\|\mathrm{H}(s - s')\|), \quad (4)$$

where H is a linear transformation representing geometric anisotropy and can be derived from two parameters (see Thorson et al. 2015, Appendix A for details), $s - s'$ is the difference in eastings and northings between locations s and s', and $\|\mathrm{H}(s - s')\|$ is the distance between

150 locations after accounting for geometric anisotropy (see Cressie and Wikle 2011, Eq. 4.9 for
151 details).

152 A piecewise constant approximation, which consists of reducing a random field w
153 defined over a spatial domain Ω to a set of knots, was used to approximate the random field w
154 (Thorson et al. 2015). To accomplish this, a desired number of knots n_j need to be pre-specified.
155 Each knot is associated with a constant value of w and covariates, and the value of w at a given
156 location s_i is determined from the value w at the knot that is nearest to s_i. A *k*-means algorithm
157 was applied to the location of survey data to determine the locations of all knots. The derived
158 distribution of knots reflects the sampling intensity of survey locations and stays the same among
159 years. The area a_j of each knot *j* was then calculated using the Voronoi tool in the *PBSmapping*
160 package in R (Schnute et al. 2013). The number of pre-specified knots is a compromise between
161 accuracy of the piecewise constant approximation and computational speed, and we confirmed
162 that all results are invariant to small increases in the number of knots used.

163 The probability of occupied habitat was modeled as a combination of linear predictors
164 (including random fields):

165 $$p_i = \text{logit}^{-1}\left(d^{(p)}_{T_{(i)}} + \sum_{k=1}^{n_x}\beta^{(p)}_k x_{J_{(i),k}} + \omega^{(p)}_{J_{(i)}} + \varepsilon^{(p)}_{J_{(i)},T_{(i)}}\right), \quad (5)$$

166 where p_i is the probability of occupied habitat for sample *i* at location s_i, $d_{T_{(i)}}$ is the average
167 density in year *t*, β_k is the coefficient of covariate x_k, J_i is the nearest knot to sample *i*, ω_i is the
168 value of random field at knot *j* that is persistent among years, $\varepsilon_{j,t}$ is the value of random field at
169 knot *j* in year *t*, and n_x is the number of covariates that are included in the model. The two
170 random fields were specified as:

171 $$\omega^{(p)} \sim MN\left(0, \Sigma^{(p)}_{\omega}\right) \quad (6)$$

$$\varepsilon_t^{(p)} \sim MN\left(0, \Sigma_{\varepsilon_t}^{(p)}\right) \quad (7)$$

Similarly, the expected positive catches λ given occupied habitat for sample i (i.e., the second model component) was specified as:

$$\lambda_i = w_i \exp\left(d_{T_{(i)}}^{(\lambda)} + \sum_{k=1}^{n_x} \beta_k^{(\lambda)} x_{J_{(i)},k} + \omega_{J_{(i)}}^{(\lambda)} + \varepsilon_{J_{(i)},T_{(i)}}^{(\lambda)}\right), \quad (8)$$

where w_i is the area swept for sample i. Random fields for positive catches are defined in the same way as probability of occupied habitat.

The total abundance across the entire modeled spatial domain can then be calculated by summing up the total abundance of all the knots:

$$\hat{b}_t = \sum_{j=1}^{n_j} a_j \operatorname{logit}^{-1}\left(\hat{d}_t^{(p)} + \sum_{k=1}^{n_x} \hat{\beta}_k^{(p)} x_{j,k} + \hat{\omega}_j^{(p)} + \hat{\varepsilon}_{j,t}^{(p)}\right)_i \exp$$

$$\left(\hat{d}_t^{(\lambda)} + \sum_{k=1}^{n_x} \hat{\beta}_k^{(\lambda)} x_{j,k} + \hat{\omega}_j^{(\lambda)} + \hat{\varepsilon}_{j,t}^{(\lambda)}\right), \quad (9)$$

where $\hat{b}_t$ is the total abundance in year t, $\hat{d}_t^{(p)}$, $\hat{\beta}_k^{(p)}$, $\hat{d}_t^{(\lambda)}$, and $\hat{\beta}_k^{(\lambda)}$ are fixed effects in the model and $\hat{\omega}_j^{(p)}$, $\hat{\varepsilon}_{j,t}^{(p)}$, $\hat{\omega}_j^{(\lambda)}$, and $\hat{\varepsilon}_{j,t}^{(\lambda)}$ are random effects in the model.

Application to northern shrimp

Northern shrimp in the GOM are at the southern extent of their range, concentrated in the southwestern region of the Gulf (Haynes and Wigley 1969; Clark et al. 1999). They are protandric hermaphrodites, maturing first as males and then transforming to females (Berkely 1931; Bergstrom 2000). In the GOM, northern shrimp are most frequently found in depths less than 300m (Haynes and Wigley 1969), with juveniles and immature males inhabiting shallower, inshore waters and adults occupying deeper offshore waters (Apollonio and Dunton 1969; Haynes and Wigley 1969; Apollonio et al. 1986). Factors that might influence shrimp

193 distribution include water temperature, salinity, depth, and substrate type (Haynes and Wigley
194 1969; Shumway et al. 1985; Apollonio et al. 1986).
195 We applied the spatio-temporal generalized linear mixed model to data collected from a
196 summer shrimp bottom trawl survey, which is designed specifically for monitoring northern
197 shrimp in the western GOM, operated with consistent sampling protocol by the Northeast
198 Fisheries Science Center in cooperation with the Atlantic States Marine Fisheries Commission
199 from 1984 to 2013. A stratified random sampling design is used to select stations sampled during
200 the survey. The survey area is divided into 12 strata and stratification is based primarily on depth,
201 latitude/longitude, and historical fishing patterns (Figure 1; Clark 1989). However, additional
202 fixed stations are also visited each year. Design-based indices of abundance and biomass for
203 stock assessment are derived from data collected at the random stations within six strata (i.e.,
204 strata 1, 3, 5, 6, 7, and 8; Figure 1) that have been sampled most intensively and consistently
205 over time. Extreme fluctuations have been observed in the design-based survey indices in recent
206 years, including a spike in the 2006 abundance estimate (see NEFSC 2014, Figure C5. 12.). Such
207 high variability could not be explained by the stock assessment models, which was one of the
208 reasons that the recent assessment was not successful for northern shrimp in the GOM (NEFSC
209 2014). In this study, we included all the data from the summer survey including non-random
210 stations in the spatio-temporal model assuming that the process of selecting sampling locations is
211 independent of the process generating differences in population density (Diggle et al. 2010).
212 We overlaid a 2 km × 2km grid on the entire survey spatial domain, which resulted in
213 4977 grid cells. For each cell we extracted the centroid and recorded the corresponding eastings
214 and northings. The value of the random field in each cell was assumed to be equal to its value at
215 the nearest knot according to the piecewise constant approximation. The value of covariates for a

given knot is the average value of the covariates for all grids that are closest to the knot. The area of a given knot can be calculated as the summation of areas of all the grid cells associated with it.

Habitat covariates and model selection

We included two static habitat variables (depth and sediment grain size), and two dynamic habitat variables (bottom temperature and bottom salinity) to estimate the occurrence and positive model. We assume that these static and dynamic habitat variables are the primary drivers of spatial variation of shrimp distribution (Shumway et al. 1985; Apollonio et al. 1986; Clark et al. 2000). During spring through autumn, adult shrimp are distributed primarily in depths between 90m and 180m (Clark et al. 2000). Temperature may impose restrictions on the amount of available habitat for northern shrimp in the GOM as seasonal bottom water temperatures in some areas can exceed the preferred range (0-5^{o}C, Shumway et al. 1985; Mountain and Jessen 1987). Adult shrimp are thought to seek deep basins as cold water refuges (Apollonio et al. 1986), therefore depth is likely to explain some variation in shrimp spatial distribution. Northern shrimp prefer an organic-rich muddy bottom (Hjort and Ruud 1938; Bigelow and Schroeder 1939; Wigley 1960; Haynes and Wigley 1969), but they are not limited to this habitat (Schick 1991). Depth and bottom temperature were recorded at each station in the survey; however, we used habitat data which are available for the entire spatial domain from other sources. Depth data were from U. S. Geological Survey, Coastal and Marine Geology Program (http://pubs.usgs.gov/of/1998/of98-801/bathy/data.htm). Sediment grain size was obtained from U. S. Geological Survey Open-File Report 2005-1001 (http://woodshole.er.usgs.gov/openfile/of2005-1001/htmldocs/datacatalog.htm#surficial_

238 sediment). The Finite-Volume Community Ocean Model (FVCOM) was used to produce bottom
239 temperature and bottom salinity data (Chen et al. 2006).

240 We only considered the habitat covariates as main effects and did not consider any
241 interactions among the covariates. For dynamic habitat variables, i.e., bottom temperature and
242 bottom salinity, effects on the response variable were assumed to be constant across years. Each
243 of the covariates was standardized to have mean of zero and unit variance prior to inclusion in
244 the model. This facilitates the interpretation of their coefficients via comparison with others in
245 the model. The value of covariates in each knot was calculated as the average values in the area
246 associated with the knot. Models with and without covariates were compared to evaluate if
247 inclusion of covariates would improve the model fit and provide more accurate and precise
248 estimates of abundance. This also allowed us to identify which habitat variable has the largest
249 influence over shrimp spatial density. We looked at whether inclusion of habitat variables
250 decreased the marginal standard deviation (MSD) of spatial and spatio-temporal variables (see
251 Thorson et al. 2015, Appendix A for details of calculating MSD). We also calculated the pseudo-
252 R^2 to determine the proportion of variance from the null model (i.e., the model has no habitat
253 variables) that was explained by including habitat variables. To do so, we compared the sum of
254 spatial and spatio-temporal variance from the null model with the same value from a model that
255 included habitat variables, and calculated the reduction in variance (pseudo-R^2) as:

256 $$pseudo\ R^2 = 1 - \frac{\sigma^2_{\omega,m}+\sigma^2_{\varepsilon,m}}{\sigma^2_{\omega,null}+\sigma^2_{\varepsilon,null}}, \quad (10)$$

257 where subscript *m* and *null* indicate a particular model *m* and the null model, respectively.

258

259 ***Model estimation***

The model includes both fixed effects (i.e., year and habitat covariates) and random effects (i.e., random fields). Fixed-effect parameters were estimated by identifying their values that maximized the marginal likelihood function. We used Template Model Builder (TMB, Kristensen et al. 2016), which approximates the marginal likelihood using the Laplace approximation and then calculates the gradient of the marginal likelihood with respect to all fixed effects. The probability of random fields was approximated using the stochastic partial differential equation approach (Lindgren et al. 2011), as explained in detail in Thorson et al. (2015). The marginal likelihood was maximized using conventional gradient-based non-linear optimization in the R statistical platform (R Core Development Team 2013). The bias of derived quantities (e.g., $\hat{b}_t$) caused by transforming nonlinear function of fixed and random effects was accounted for by using a newly developed bias-correction algorithm. Further details can be found in Thorson and Kristensen (2016). We used R package SpatialDeltaGLMM to estimate all the parameters of the spatio-temporal index standardization model (https://github.com/nwfsc-assess/geostatistical_delta-GLMM).

Comparison between model-based and design-based indices

We compared the performance of a stock assessment model for northern shrimp when fit with spatio-temporal indices and design-based indices (stratified mean). The model was a seasonal size-structured assessment model developed for hermaphroditic Pandalidae (Cao et al. 2016) and first vetted in a northern shrimp benchmark stock assessment (NEFSC 2014). The spatio-temporal index used to fit the assessment model was derived from the estimated densities restricted to the same six strata used to estimate the design-based index. We compared model fit and retrospective patterns for spawning stock biomass and recruitment produced by the

283 assessment model using both spatio-temporal and design-based indices. Retrospective bias was
284 quantified using a revised rho statistic of Mohn (1999) (Hurtado-Ferro et al. 2014). We also
285 evaluated the discrepancy between model outputs (i.e., spawning stock biomass and recruitment)
286 and biological reference points estimated based on spatio-temporal and design-based indices.
287 Spawning potential ratio-based metrics were calculated to determine the stock status for both
288 assessments using spatio-temporal and design-based indices.

289 Finally, we conducted cross-validation tests using a retrospective method for evaluating
290 the predictive performance of the size-structured assessment model using design-based and
291 spatio-temporal indices. One-year-ahead abundance index was forecasted based on the estimates
292 from the stock assessment with a series of fits in which y = 1,…,13 years of data are left out at
293 the end. Relative error was calculated to quantify the disparity between model forecast and the
294 observed index for a given year:

295 $$RE_y = \frac{I^{pred}_{2014-y} - I^{obs}_{2014-y}}{I^{obs}_{2014-y}} \quad (11)$$

296 where I^{pred}_{2014-y} is the median forecasted value in log-space for year 2014-y and I^{obs}_{2014-y} is the
297 design-based or spatio-temporal index in log-space for year 2014-y. Uncertainty in the forecasted
298 abundance index arises from uncertainty in estimated recruitment of the forecasting year.
299 Average absolute relative error was calculated for comparing the model predictive performance
300 with design-based and spatio-temporal indices:

301 $$ARE = \frac{\sum_{y=1}^{13} |RE_y|}{13} \quad (12)$$

302

303 **RESULTS**

304 The estimated MSD was very close to zero for spatial and spatio-temporal variation in
305 "true" zeros (i.e., $\omega^{(p)}$ and $\varepsilon^{(p)}$). Therefore, we turned off the random fields for modeling the
306 probability of occupied habitat (i.e., $\omega^{(p)}$ and $\varepsilon^{(p)}$). Including habitat variables did not decrease
307 the MSD of spatio-temporal variation ($\sigma_{\varepsilon}^{(\lambda)}$). However, it slightly reduced the MSD of spatial
308 variation $\sigma_{\omega}^{(\lambda)}$ (Table 1). In combination, habitat variables explained 25.9% of spatial and spatio-
309 temporal variation, which suggests that these habitat variables have an important impact on (but
310 do not by themselves fully explain) spatial variation in shrimp density. In general, spatial
311 variation that is constant over time has a greater magnitude than spatial variation that changes
312 annually. Therefore, we identified the model included depth, sediment grain size and salinity as
313 covariates, which had the lowest MSD for spatial variation and highest pseudo-R^2 (0.262), as the
314 base model for further analysis.
315 Comparison of the MSDs of spatial and spatio-temporal variation estimated from the
316 model with and without each habitat variable indicates that the largest portion of the decreased
317 spatial variation for shrimp densities was explained by sediment grain size (Table 1), so we
318 concluded that this variable had the most significant impact on shrimp densities. Salinity, by
319 itself, did not lead to decreased spatial variation but made an important contribution in the
320 presence of other habitat variables. Temperature, alone or in the presence of other variables, did
321 not contribute to the decreased spatial variation for shrimp densities.
322 The spatio-temporal variation in shrimp density is shown in Figure 2, estimated from the
323 base model. The highest densities were generally found in the vicinity of Jeffreys Ledge, while
324 the lowest densities were found in the southeast of Cape Cod and southeast boundary of the
325 survey area (Figure 2). However, shrimp densities dropped dramatically in the recent two years,
326 especially for the southern portion of the survey domain where almost no shrimp could be found.

The model's raw residuals did not show a strong spatial pattern over years (Figure 3). However, the model fits to data after 2005 were worse than the fits to data from years prior. The estimated anisotropy for the model component for positive catches showed that spatial residuals in positive catches were stretched along the Northeast – Southwest, suggesting that densities are correlated over a longer distance moving along the shoreline than perpendicular to the shoreline.

We next compared the estimated spatio-temporal and design-based indices. Given that the assessment model freely estimates the catchability coefficient associated with this index, the only information it provides is regarding relative trends. We therefore standardize all indices prior to comparing them. The design-based index showed greater temporal variation than the spatio-temporal index. The spatio-temporal indices for 2006 were much lower than the corresponding design-based index (spatio-temporal indices: 3.92; design-based index: 6.86). However, the design-based and spatio-temporal indices showed similar temporal trends, except some minor discrepancies during the early time period (Figure 4). Trends in spatio-temporal indices calculated based on six strata and the entire area were almost identical (Figure 4).

We then fit the size-structured assessment model using the estimated spatio-temporal index (based on six strata) and compared the results with those obtained from the model using the design-based index. The results showed improvement in the assessment model fit overall (total likelihood decreased from 9971.02 to 9945.33). More specifically, there was a substantial decrease in the likelihood of the abundance index (from 17.993 to -6.795). The predicted abundance index from the assessment model was within or very close to the 95% interval of the centered spatio-temporal index for every year (Figure 5a). However, the assessment model prediction disagreed with the design-based index for some years, e.g., 1994, 2006 and 2007 (Figure 5b). More importantly, the assessment model based on the design-based index failed to

capture the spike observed in the design-based index of 2006 (Figure 5b). However, the assessment model using the spatio-temporal index captured this variation reasonably well (Figure 5a). This suggests that the index derived from the spatio-temporal model provided more consistent information with other assessment data inputs (e.g., total catch, catch and survey compositions). Improvement in model diagnostics when using the spatio-temporal index was also found in retrospective patterns. Retrospective pattern was reduced when the assessment model was fitted to the spatio-temporal index (Figure 6). Revised Mohn's rho, measured for estimated spawning stock biomass and recruitment, reduced by 66.7% and 20%, respectively, when the spatio-temporal index was used in the assessment model.

We also compared the outputs from assessment models (i.e., estimated recruitment and spawning stock biomass) using the spatio-temporal and design-based indices. Both recruitment and spawning stock biomass estimates based on the spatio-temporal index were more than 50% larger than those based on design-based index for the recent years, i.e., 2010 – 2013 (Figure 7), which are of high importance to managers. On average, spawning stock biomass was more sensitive to the changes in abundance index than recruitment, and spawning stock biomass estimates based on the design-based index were larger than those based on the spatio-temporal index on average. Proxy reference points based on spawning potential ratio were calculated to determine the stock status. $F_{40\%}$ based on spatio-temporal and design-based indices were 0.78 and 0.83, respectively. Forty percent of unexploited spawning stock biomass based on spatio-temporal and design-based indices were 2375.9 and 2433.3 mt, respectively. Historical stock status determined based on design-based index was more optimistic than that based on spatio-temporal index (Figure 8).

Finally, we evaluated the predictive performance of assessment models (i.e., forecasting one-year-ahead abundance index) using spatio-temporal and design-based indices. On average, the assessment model using the spatio-temporal index performed better in forecasting the one-year-ahead abundance index (*ARE* for spatio-temporal index: 0.52; *ARE* for design-based index: 0.62) (Figure 9). The assessment model using the design-based index did not converge when forecasting the abundance index for years 2005, 2010, 2012 and 2013. However, the assessment model using the spatio-temporal index successfully forecasted the abundance indices for all the 13 years. The models showed similar performance in forecasting abundance indices of 2002, 2008 and 2011, but for 2001, 2004 and 2009 the model using the design-based index had relative error closer to zero than the model using spatio-temporal index (Figure 9). However, the assessment model using the spatio-temporal index greatly outperformed the model using the design-based index in forecasting abundance indices of 2006 and 2007 when the abundance indices showed large temporal variation.

DISCUSSION

In this study, we evaluated a potential improvement to a size-structured assessment model for northern shrimp in the GOM. We showed that using a spatio-temporal abundance index (rather than an abundance index derived from classical design-based estimators) resulted in improved predictive performance (from a one-step-ahead predictive evaluation) and retrospective performance (using Mohn's rho). We therefore conclude that using the spatio-temporal index resulted in improvements in assessment model performance for this stock. The spatio-temporal and design-based estimators resulted in particularly large differences in 2006, but otherwise had similar abundance trends. Despite these similarities in abundance trends,

utilizing the spatio-temporal index in the assessment model greatly alters the estimates of recruitment and spawning stock biomass, especially for the recent years (Figure 7), and also alters the determination of stock status for some years (Figure 8). Based on the spawning potential ratio-based metrics, the assessment using a design-based estimator leads to a more optimistic perception of historical stock status for northern shrimp in the GOM (Figure 8).

Spatially correlated variation in density is observed in almost all fisheries data collected from both fishery-dependent and fishery-independent sources (Booth 2000). However, spatial variation is often ignored or not properly dealt with in statistical analysis and inference. Consequently, it results in inaccurate and imprecise estimates of relative abundance (Swartzman et al. 1992; Petitgas 1993) and/or misleading interpretations of various aspects of a species' biology (Thorson 2015). By contrast, the spatio-temporal index standardization can provide a more precise abundance index than design-based estimator or conventional models by explaining spatial variation in densities (Shelton et al. 2014; Thorson et al. 2015). Specifically, densities in different locations are assumed to have distinct expected values based on habitat covariates and spatial terms, and densities at nearby sites are more similar than densities at geographically remote sites. In contrast, design-based estimator assumes that the mean of a given stratum is fixed and all locations within that stratum provide exchangeable samples of a single mean. Thus, the design-based estimator is often more sensitive to outlier observations (Shelton et al. 2014). This effect can be particularly significant in conjunction with decreasing sampling effort. This might be the reason that the centered design-based index of 2006 is much larger than corresponding spatio-temporal index in this study. The number of sampling locations used to derive the design-based index in 2006 (29) is smaller than other years (40 – 50), and 5 out of 29 sampling locations fell in 'hot spots' which had one order of magnitude higher tows than average.

In this case, relative abundance estimates can benefit greatly from filling spatial gaps (i.e., predicting un-sampled locations) using a spatio-temporal model.

We included habitat covariates in the spatio-temporal model to capture the important structure in the mean. This could partially account for the non-stationarity which could arise from two sources. However, we did not use non-stationary Gaussian random fields in our study because a previous study suggested that non-stationary Gaussian random fields are not always necessary to model non-stationary spatial data (Fuglstad et al. 2015). We have shown that including depth, sediment grain size and salinity as covariates in the model explained approximately 26.2% percent of spatial and spatio-temporal variation relative to the model that had no habitat variables, and did so by decreasing the spatial (constant over time) component. Northern shrimp prefer fine-grained sediments according to our study, perhaps because they provide more food, e.g., soft bottom benthic invertebrates. Depth had a positive effect on shrimp density as deeper basins are thought to provide cold water refuges (Apollonio et al. 1986). These results are consistent with previous studies (Haynes and Wigley 1969; Shumway et al. 1985). The range of salinity fields used in the model was relatively narrow (31.3 – 35.0 psu), which is well within the salinity range reported in the previous study (Shumway et al. 1985). Within that range of salinity, our study indicated that shrimp prefer slightly lower salinity in the study area. However, bottom temperature did not contribute to explaining spatial variation in our study, despite its well-documented influence on shrimp population dynamics in this region (Apollonio et al. 1986; Richards et al. 2012; Richards et al. 2016). Summer bottom temperatures in shrimp habitat areas have remained several degrees cooler than upper thermal tolerance levels for adult northern shrimp (Shumway et al. 1985; Bergstrom 2000), even with warming in recent years. Thus the thermal gradient may have been too weak during the study period to influence the

441 distribution of shrimp. Also, it is possible that the effect of temperature on shrimp density is at
442 much finer spatial scale which could not be well approximated by the knots. To evaluate this
443 possibility, we calculated the percentage of total variance in the habitat variables that was
444 explained by variance among knots. The results showed that except for depth (72%) the
445 percentages for other variables are about 50%, suggesting that fine-scale variability of habitat
446 covariates was not used in the model.

447 Additionally, our spatio-temporal model, which can be considered as an extension of
448 generalized linear model, assumes that the relationship between habitat and response variables is
449 linear. Thus, the model would need to be modified to account for dome-shaped or saturating
450 relationships between habitat variables and shrimp density. We avoided adding polynomial
451 expansion x^2, x^3 and interaction terms to the model because of the risk of overfitting and the
452 extensive cross-validation testing required to avoid overfitting. Using a habitat suitability index
453 as the only habitat covariate in the spatio-temporal model might better explain the spatial
454 variation while keeping the model parsimonious (Breece et al. 2016).

455 We also note that sampling intensity in marginal strata changed over time as a result of
456 preferential sampling that led to the over-sampling locations corresponding to high densities (i.e.,
457 strata used to derive design-based index). The standard deviations of predicted densities were
458 high in the area where sampling intensity was low (Figure S1). Therefore, we cannot eliminate
459 the possibility that model-based inference for the entire survey domain is biased. However, we
460 used the spatio-temporal index derived from six strata for comparison, so the conclusions of this
461 study are less likely to be influenced by violating the model assumption of non-preferential
462 sampling. We suggest that future research could explore the spatio-temporal models for
463 preferential sampling (Diggle et al. 2010). The spatio-temporal abundance index was estimated

464 from the spatio-temporal model based on the data including non-random stations. However,
465 excluding the data from non-random stations did not appreciably change the abundance estimates
466 (Figure S2).

467 We envision several important topics for future applications of spatio-temporal
468 estimation methods. Most importantly, spatio-temporal methods could be used to estimate
469 density for different size or age-classes of fishes and invertebrates (Kristensen et al. 2014). These
470 estimates could then be processed to generate age or size-composition data for assessment
471 models. A model-based approach to estimating age- or size-composition may be more
472 statistically efficient for species with spatial segregation of size or age groups (e.g. life history
473 stages). This spatial predictability is not currently used by design-based or stratified approaches
474 to compositional standardization (Thorson 2014). For example, if male northern shrimp are
475 preferentially distributed in shallow waters, then we expect that design-based estimates of size
476 composition would be skewed towards male due to the preferential random allocation of sample
477 locations (i.e., sampled most intensively in strata 1, 3, 5, 6, 7, and 8). Compositional data have a
478 strong effect on assessment results for many species (for better or worse; Francis 2011), so we
479 highly recommend methods to improve statistical efficiency for these data.

480 We also recommend continuing research to improve statistical efficiency when
481 estimating abundance trends from survey samples. We note that multispecies data are a generally
482 under-utilized source of information regarding habitat suitability. In particular, detecting a
483 species with similar habitat preferences may be informative about the likely density of a target
484 species (Thorson et al. 2015). We therefore suspect that jointly analyzing survey catch rates for
485 multiple species may improve density estimates for rare or poorly-sampled species. Planned
486 surveys following a randomized design continue to be the most reliable source of information

regarding stock status for fisheries worldwide. We therefore encourage any research that allows better inference to be made using limited historical and expensive ongoing surveys.

ACKNOWLEDEGEMENTS

Financial support for this study was provided by the Maine Sea Grant Program, the Maine Department of Marine Resources and the NOAA FATE program. We thank Margaret Hunter and the Atlantic States Marine Fisheries Commission Northern Shrimp Technical Committee for informative discussions on the northern shrimp stock assessment and for providing the data on which this study is based.

REFERENCES

Apollonio, S., and Dunton, E.E. 1969. The northern shrimp, *Pandalus borealis*, in the Gulf of Maine. Completion Rept., ME Dept. Sea and Shore Fisheries, Proj. 3–12–R, 81p.

Apollonio, S., Stevenson, D., and Dunton, E.E. 1986. Effects of temperature on the biology of the northern shrimp, *Pandalus borealis*, in the Gulf of Maine. NOAA Tech. Rep. NMFS 42. Available online at: http://spo.nwr.noaa.gov/tr42.pdf

Bigelow, H.A., and Schroeder, W.C. 1939. Notes on the fauna above mud bottoms in deep water in the Gulf of Maine. Biol. Bull. (Woods Hole) 76: 305–324.

Booth, A.J. 2000. Incorporating the spatial component of fisheries data into stock assessment models. ICES J. Mar. Sci. 57(4): 858–865.

Breece, M.W., Fox, D.A., Dunton, K.J., Frisk, M.G., Jordaan, A. and Oliver, M.J. 2016. Dynamic seascapes predict the marine occurrence of an endangered species: Atlantic Sturgeon *Acipenser oxyrinchus oxyrinchus*. Methods Ecol. Evol. 7: 725–733.

Cao, J., Chen, Y., Chang, J.H., and Chen, X. 2014. An evaluation of an inshore bottom trawl survey design for American lobster (*Homarus americanus*) using computer simulations. J. Northwest Atl. Fish. Sci. 46: 27–39.

Cao, J., Chen, Y., Richards, A. 2017. Improving assessment of Pandalus stocks using a seasonal, size-structured assessment model with environmental variables: Part I: Model description and application. Can. J. Fish. Aquat. Sci. 74(3): 349–362.

Chen, C., Beardsley, R.C., Cowles, G.W. 2006. An unstructured-grid, finite-volume coastal ocean model (FVCOM) system. Oceanography 19: 78–89.

518 Clark, S.H. 1989. State-Federal northern shrimp survey. In: Proceedings of a workshop on
519 bottom trawl surveys. Azarovitz, T.R., McGurrin, J. and Seagraves, R. (eds.). ASMFC Spec.
520 Rept. 17: 27–29.

521 Clark, S.H., Cadrin, S.X., Schick, D.F., Diodati, P.J., Armstrong, M.P., and McCarron, D. 2000.
522 The Gulf of Maine northern shrimp (*Pandalus borealis*) fishery: a review of the record. J.
523 Northw. Atl. Fish. Sci. 27: 193–226.

524 Cressie, N., and Wikle, C.K. 2011. Statistics for Spatio-Temporal Data. JohnWiley & Sons.

525 Diggle, P.J., Menezes, R., and Su, T. 2010. Geostatistical inference under preferential sampling.
526 J. Roy. Statist. Soc. Ser. C 59: 191–232.

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

529 Fuglstad, G.A., Simpson, D., Lindgren, F. and Rue, H. 2015. Does non-stationary spatial data
530 always require non-stationary random fields? Spat. Stat.14: 505–531.

531 Haynes, E.B., and Wigley, R.L. 1969. Biology of the northern shrimp, *Pandalus borealis*, in the
532 Gulf of Maine. Trans. Am. Fish. Soc. 98(1): 60–76.

533 Helser, T.E., Punt, A.E., and Methot, R.D. 2004. A generalized linear mixed model analysis of a
534 multi-vessel fishery resource survey. Fish. Res. 70: 251–264.

535 Hjort, J., and Ruud, J.T. 1938. Deep-sea prawn fisheries and their problems. Hvalr. Skr. 17: 1–
536 44.

537 Hurtado-Ferro, F., Szuwalski, C.S., Valero, J.L., Anderson, S.C., Cunningham, C.J., Johnson, K.
538 F., Licandeo, R., McGilliard, C.R., Monnahan, C.C., Muradian, M.L., Ono, K., Vert-Pre, K.A.,
539 Whitten, A.R., and Punt, A.E. 2014. Looking in the rear-view mirror: bias and retrospective
540 patterns in integrated, age-structured stock assessment models. ICES J. Mar. Sci. 72: 99–110.

541 Kristensen, K., Thygesen, U.H., Andersen, K.H., and Beyer, J.E. 2014. Estimating spatio-
542 temporal dynamics of size-structured populations. Can. J. Fish. Aquat. Sci. 71: 326–336.

543 Kristensen, K., Nielsen, A., Berg, C.W., Skaug, H., and Bell, B.M. 2016. TMB: Automatic
544 Differentiation and Laplace Approximation. J. Stat. Softw. 70(5): 1–21.

545 Lindgren, F., Rue, H., and Lindstrom, J. 2011. An explicit link between Gaussian fields and
546 Gaussian Markov random fields: the stochastic partial differential equation approach. J. Roy.
547 Statist. Soc. Ser. B 73: 423–498.

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

551 Maunder, M.N., and Punt, A.E. 2004. Standardizing catch and effort data: a review of recent
552 approaches. Fish. Res. 70: 141–159.

553 Mountain, D.G., and Jessen, P.F. 1987. Bottom waters of the Gulf of Maine, 1978-1983. J. Mar.
554 Res. 45: 319–345.

555 Northeast Fisheries Science Center (NEFSC). 2014. 58th Northeast Regional Stock Assessment
556 Workshop (58th SAW) Assessment Report. US Dept Commer, Northeast Fish Sci Cent Ref Doc.

14-04; 784 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, or online at http://nefsc.noaa.gov/publications/

Petitgas, P. 1993. Geostatistics for fish stock assessments: a review and an acoustic application. ICES J. Mar. Sci. 50(3): 285–298.

R Core Development Team. 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Richards, R. A., Fogarty, M., Mountain, D., Taylor, M. 2012. Climate change and northern shrimp recruitment variability in the Gulf of Maine. Mar. Ecol. Prog. Ser. 464: 167–178.

Richards, R.A., O'Reilly, J. E., Hyde, K.W.J. 2016. Use of satellite data to identify critical periods for early life survival of northern shrimp in the Gulf of Maine. Fish. Oceanogr. 25(3): 306–319.

Schick, D.F. 1991. Pandalid shrimp distribution relative to bottom type and availability to commercial and research trawls in the Gulf of Maine. ICES C.M. 1991/K:8 p. 7.

Schnute, J.T., Boers, N., Haigh, R., Grandin, C., Johnson, A., Wessel, P., and Antonio, F. 2013. PBSmapping: Mapping Fisheries Data and Spatial Analysis Tools. http://CRAN.R-project.org/package=PBSmapping

Shelton, A.O., Thorson, J.T., Ward, E.J., and Feist, B.E. 2014. Spatial semiparametric models improve estimates of species abundance and distribution. Can. J. Fish. Aquat. Sci. 71: 1655–1666.

576 Shumway, S.E., Perkins, H.C., Schick, D.F., and Stickney, A.P. 1985. Synopsis of the biological
577 data of the pink shrimp *Pandalus borealis* (Krøyer, 1938) FAO Fisheries Synopsis, No 144;
578 NOAA Tech. Rep., NMFS 30, 57p. Available online at: http://spo.nwr.noaa.gov/tr30.pdf

579 Skaug, H., and Fournier, D. 2006. Automatic approximation of the marginal likelihood in non-
580 Gaussian hierarchical models. Comput. Stat. Data An. 51: 699–709.

581 Smith, S. J. 1990. Use of statistical models for the estimation of abundance from groundfish
582 trawl survey data. Can. J. Fish. Aquat. Sci. 47: 894–903.

583 Swartzman, G., Huang, C., and Kaluzny, S. 1992. Spatial analysis of Bering Sea groundfish
584 survey data using Generalized Additive Models. Can. J. Fish. Aquat. Sci. 49: 1366–1378.

585 Thorson, J.T. 2014. Standardizing compositional data for stock assessment. ICES J. Mar. Sci. 71:
586 1117–1128.

587 Thorson, J.T. 2015. Spatio-temporal variation in fish condition is not consistently explained by
588 density, temperature, or season for California Current groundfishes. Mar. Ecol. Prog. Ser. 526:
589 101–112.

590 Thorson, J.T., and Kristensen, K. 2016. Implementing a generic method for bias correction in
591 statistical models using random effects, with spatial and population dynamics examples. Fish.
592 Res. 175: 66–74.

593 Thorson, J.T., Scheuerell, M.D., Shelton, A.O., See, K.E., Skaug, H.J., Kristensen, K., 2015.
594 Spatial factor analysis: a new tool for estimating joint species distributions and correlations in
595 species range. Methods Ecol. Evol. 6: 627–637.

596 Thorson, J.T., Shelton, A.O., Ward, E.J., and Skaug, H.J. 2015. Geostatistical delta-generalized
597 linear mixed models improve precision for estimated abundance indices for West Coast
598 groundfishes. ICES J. Mar. Sci. 72(5): 1297–1310.

599 Thorson, J.T., and Ward, E. 2013. Accounting for space-time interactions in index
600 standardization models. Fish. Res. 147: 426–433.

601 Thorson, J.T., and Ward, E.J. 2014. Accounting for vessel effects when standardizing catch rates
602 from cooperative surveys. Fish. Res. 155: 168–176.

603 Wigley, R.L. 1960. Note on the distribution of Pandalidae (Crustacea, Decapoda) in New
604 England waters. Ecol. 41: 564–570.

605 Ye, Y., and Darren, D. 2009. How reliable are the abundance indices derived from commercial
606 catch-effort standardization? Can. J. Fish. Aquat. Sci. 66(7): 1169–1178.

607 **Table 1.** Marginal standard deviation (MSD) of spatial and spatio-temporal variables and
608 pseudo-R^2 showing the proportion of variance from the null model (i.e., the model with no
609 habitat covariates included) that is explained by including covariate(s) in the model. Note that
610 model with temperature and depth as covariates could not produce converged results. The
611 saturated model includes depth, sediment, temperature, and salinity as covariates. The model in
612 boldface is the base model.

Model	**Random fields (MSD)**		**Pseudo-R^2**
	$\sigma_\omega^{(\lambda)}$	$\sigma_\varepsilon^{(\lambda)}$	
Saturated model	0.705	0.215	0.259
Saturated model - temperature	**0.704**	**0.214**	**0.262**
Saturated model - salinity	0.748	0.217	0.173
Saturated model - depth	0.716	0.217	0.237
Saturated model - sediment	0.757	0.213	0.157
Null model + depth + sediment	0.752	0.215	0.166
Null model + temperature + salinity	0.818	0.211	0.027
Null model + depth + salinity	0.760	0.212	0.151
Null model + temperature + depth	-	-	-
Null model + temperature + sediment	0.742	0.217	0.185
Null model + sediment + salinity	0.715	0.215	0.240
Null model + depth	0.803	0.210	0.061
Null model + sediment	0.748	0.215	0.174
Null model + temperature	0.830	0.210	0.001
Null model + salinity	0.823	0.209	0.017
Null model	0.831	0.207	

613 $\sigma_\omega^{(\lambda)}$ and $\sigma_\varepsilon^{(\lambda)}$ are the MSDs of spatial and spatio-temporal random fields for the expected positive
614 catches given occupied habitat, i.e., the standard deviation of different realizations of density
615 governed by the same stochastic process.

FIGURE CAPTIONS

Figure 1. Northern shrimp summer survey area and strata in the Gulf of Maine (coastlines data from R package ocedata).

Figure 2. Density of northern shrimp 1984 – 2013, estimated by the spatio-temporal generalized linear mixed model. Predicted density is shown for the centroid of each 2 km × 2km grid cell.

Figure 3. The raw residuals (1984 – 2013) from the spatio-temporal generalized linear mixed model. Positive and negative values are in red and blue circles, respectively.

Figure 4. Centered abundance indices derived from design-based and model-based spatio-temporal approaches. Design-based index is calculated based on data from six strata (i.e., strata 1, 3, 5, 6, 7, and 8; Figure 1). Two spatio-temporal indices are estimated for different spatial areas (i.e., strata 1, 3, 5, 6, 7, and 8 and all strata). Note that coefficients of variance (CV) of abundance indices derived from design-based and model-based spatio-temporal estimators (based on six strata and all strata) are 1.20, 0.69, and 0.71, respectively.

Figure 5. Comparison of stock assessment model fits to (a) spatio-temporal index and (b) design-based index. Points show predictions from stock assessment model and red lines represent estimated abundance index with 95% intervals for the (a) spatio-temporal model and (b) design-based approach.

Figure 6. Retrospective analysis of spawning stock biomass and recruitment for assessment based on design-based and spatio-temporal indices. The full assessment time series (line extending through 2013) is compared with model runs of identical structure but with 1, 2, …, 8 years of data removed (lines extending through 2005 to 2012) to illustrate retrospective bias,

638 which is quantified by Mohn's ρ (the value is zero when the peeled assessments match exactly
639 with full time series assessment).

640 **Figure 7**. Relative changes in percentage for estimated recruitment and spawning stock biomass
641 based on spatio-temporal and design-based indices. Note that the reference value is the estimates
642 based on design-based index (for values greater than the reference value, the relative change in
643 percentage should be a positive number).

644 **Figure 8**. Status of northern shrimp stock in the Gulf of Maine determined based on stock
645 assessment with (a) design-based index and (b) spatio-temporal index. The horizontal line (red
646 and yellow) represents $F_{40\%}$ (the fishing mortality at which spawning stock biomass per recruit is
647 40% of virgin level) and the area above the line indicates that overfishing is occurring. The
648 vertical line represents spawning stock biomass at 40% of virgin spawning stock biomass and the
649 area to the left indicates that the stock has been overfished.

650 **Figure 9**. Relative error of one-year-ahead forecast index based on assessment model using
651 design-based and spatio-temporal indices. Note that the assessment model using design-based
652 index fails to forecast the abundance index for years 2005, 2010, 2012 and 2013 because of non-
653 convergence.

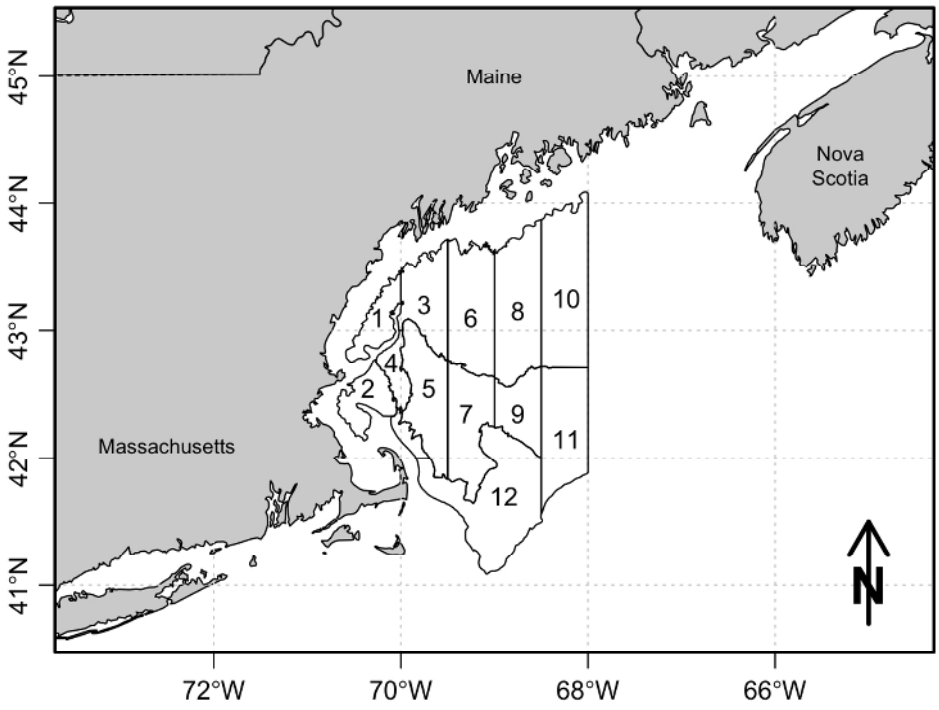

Figure 1. Northern shrimp summer survey area and strata in the Gulf of Maine (coastlines data from R package ocedata).

458x352mm (72 x 72 DPI)

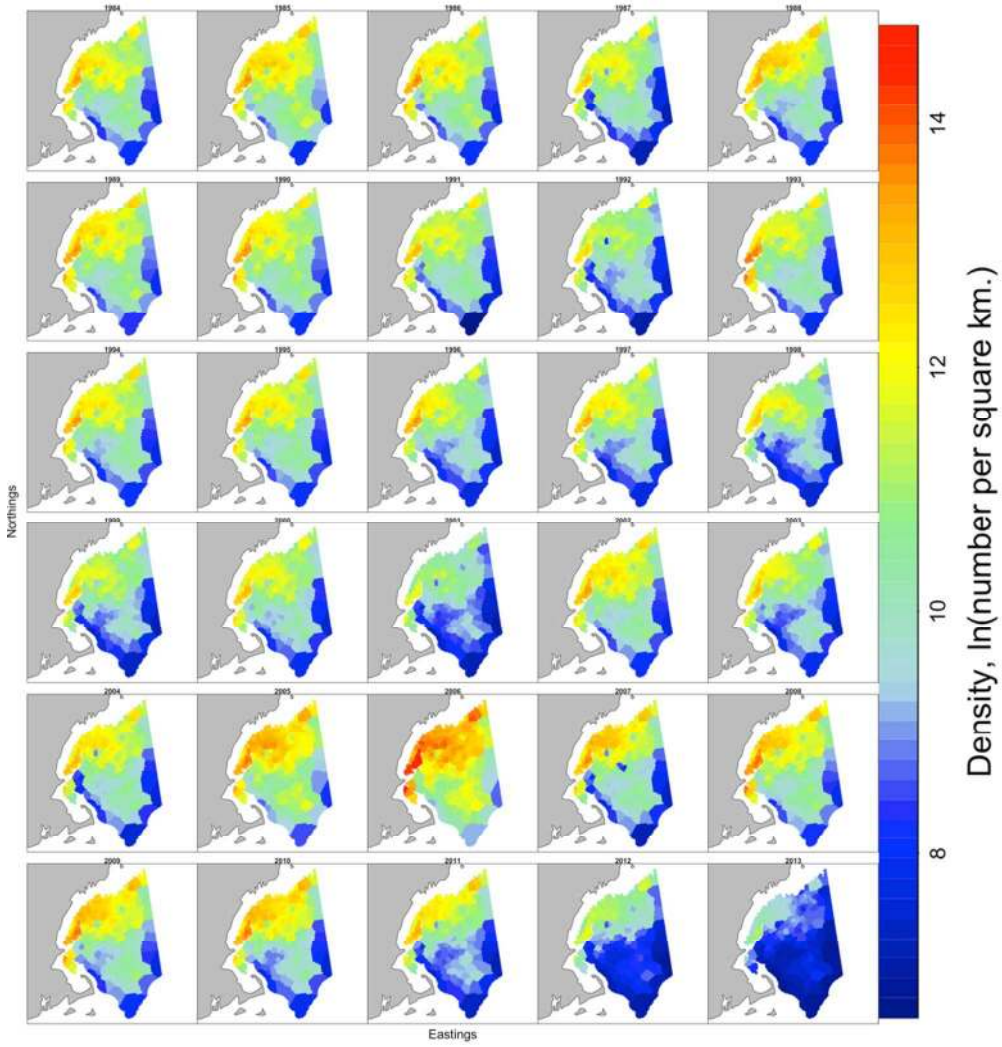

Figure 2. Density of northern shrimp 1984 – 2013, estimated by the spatio-temporal generalized linear mixed model. Predicted density is shown for the centroid of each 2 km × 2km grid cell.

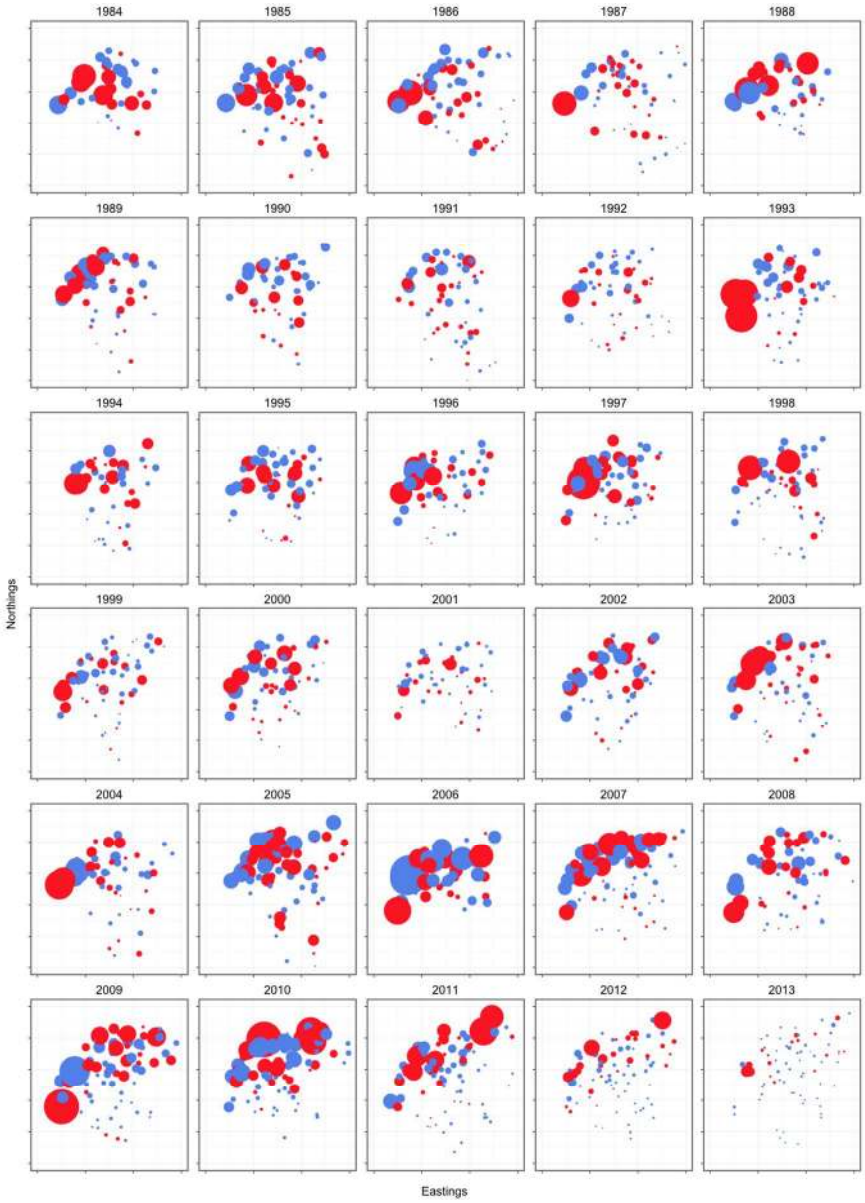

Figure 3. The raw residuals (1984 – 2013) from the spatio-temporal generalized linear mixed model. Positive and negative values are in red and blue circles, respectively.

458x635mm (72 x 72 DPI)

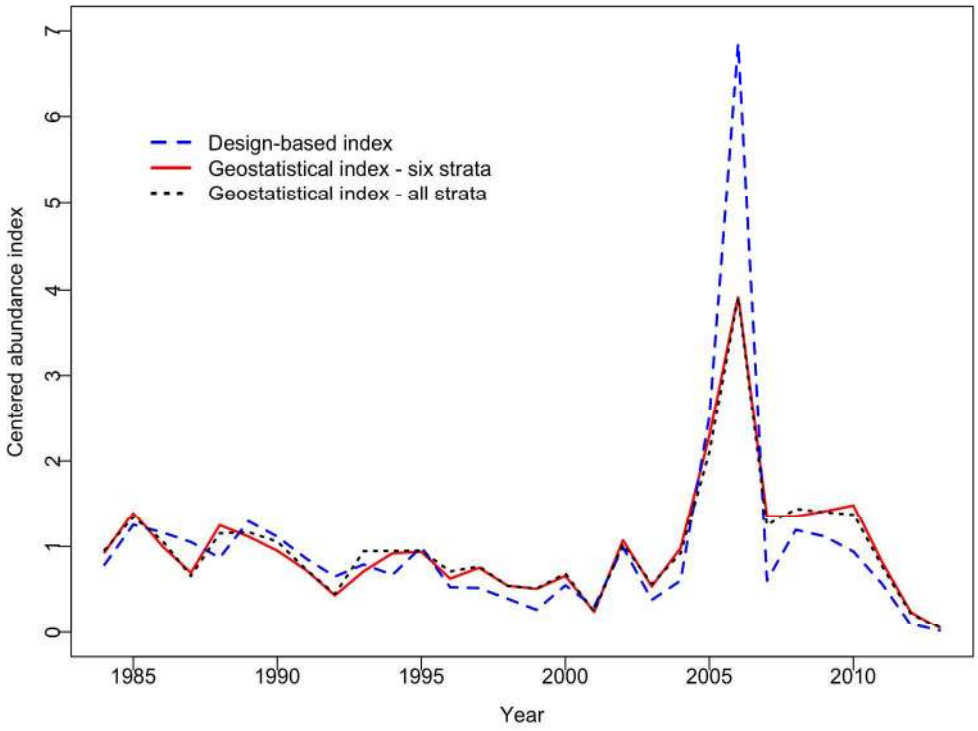

Figure 4. Centered abundance indices derived from design-based and model-based spatio-temporal approaches. Design-based index is calculated based on data from six strata (i.e., strata 1, 3, 5, 6, 7, and 8; Figure 1). Two spatio-temporal indices are estimated for different spatial areas (i.e., strata 1, 3, 5, 6, 7, and 8 and all strata). Note that coefficients of variance (CV) of abundance indices derived from design-based and model-based spatio-temporal estimators (based on six strata and all strata) are 1.20, 0.69, and 0.71, respectively.

846x635mm (72 x 72 DPI)

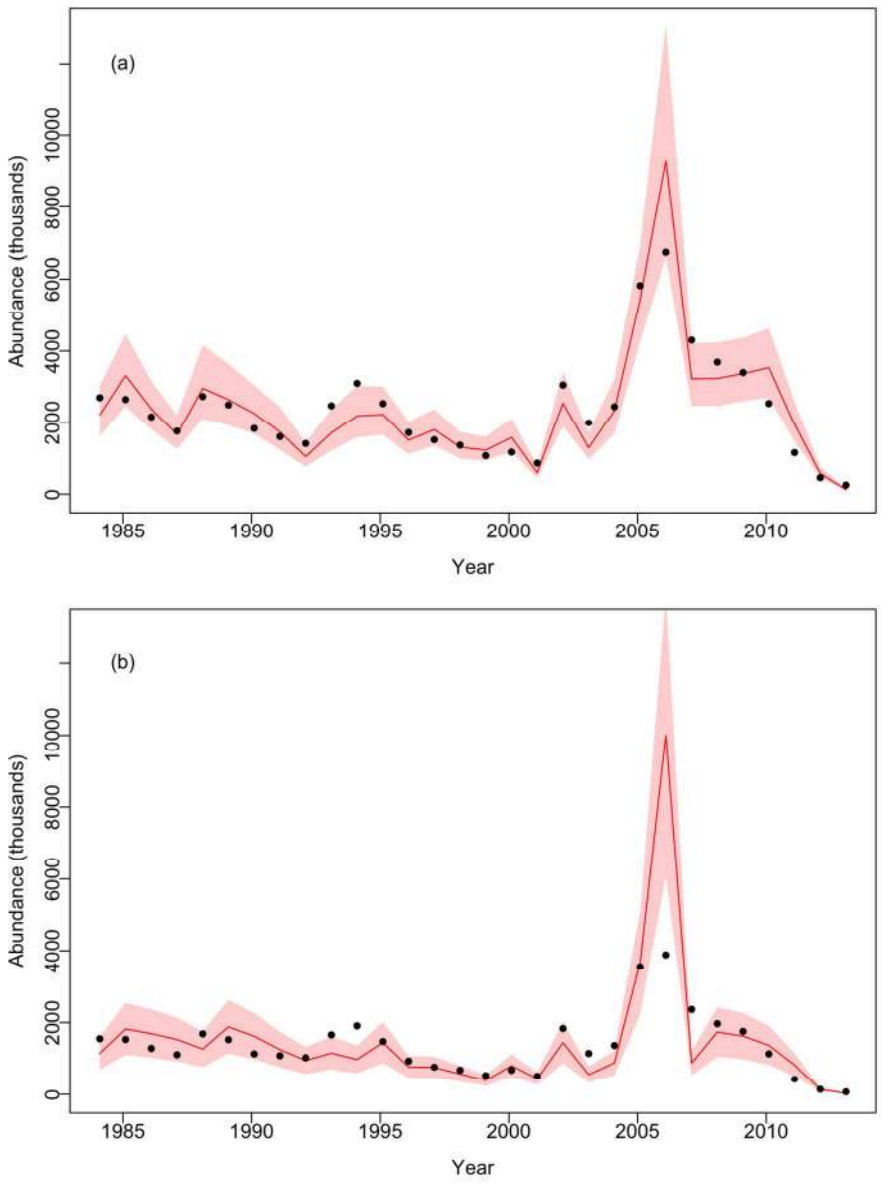

Figure 5. Comparison of stock assessment model fits to (a) spatio-temporal index and (b) design-based index. Points show predictions from stock assessment model and red lines represent estimated abundance index with 95% intervals for the (a) spatio-temporal model and (b) design-based approach.

529x705mm (72 x 72 DPI)

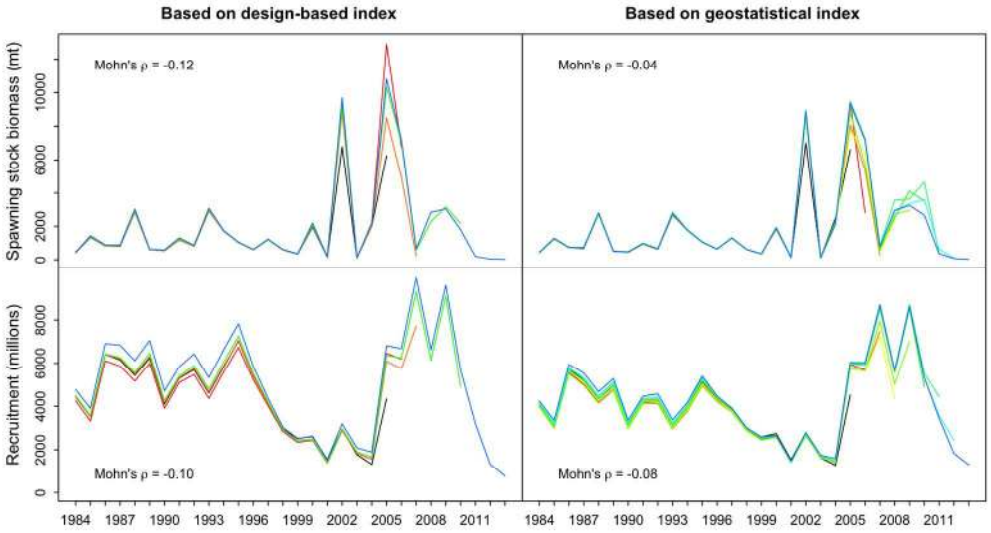

Figure 6. Retrospective analysis of spawning stock biomass and recruitment for assessment based on design-based and spatio-temporal indices. The full assessment time series (line extending through 2013) is compared with model runs of identical structure but with 1, 2, …, 8 years of data removed (lines extending through 2005 to 2012) to illustrate retrospective bias, which is quantified by Mohn's ρ (the value is zero when the peeled assessments match exactly with full time series assessment).

705x423mm (72 x 72 DPI)

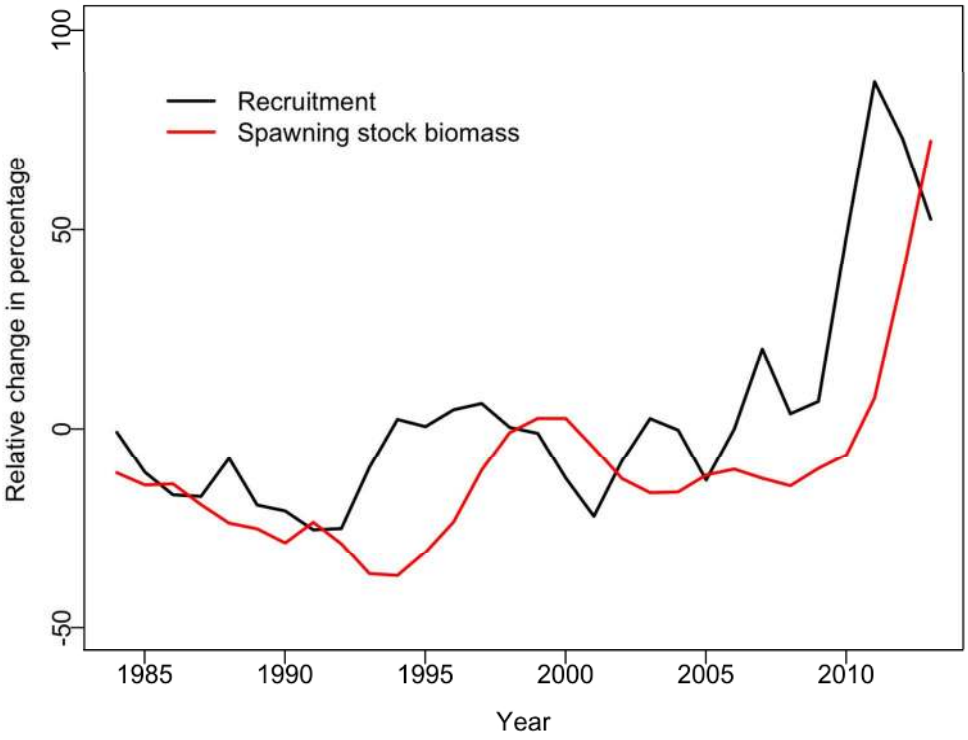

Figure 7. Relative changes in percentage for estimated recruitment and spawning stock biomass based on spatio-temporal and design-based indices. Note that the reference value is the estimates based on design-based index (for values greater than the reference value, the relative change in percentage should be a positive number).

458x352mm (72 x 72 DPI)

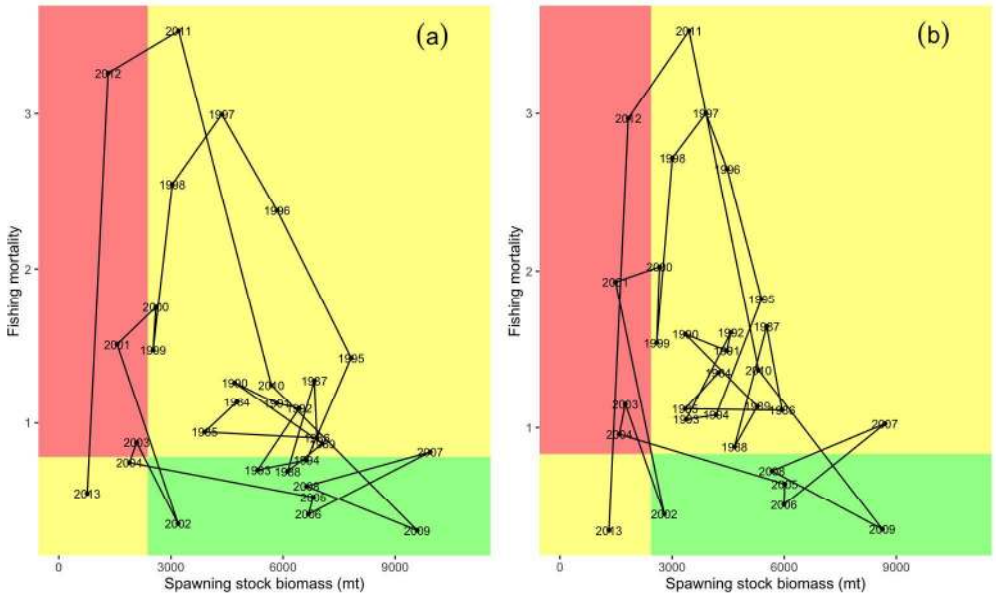

Figure 8. Status of northern shrimp stock in the Gulf of Maine determined based on stock assessment with (a) design-based index and (b) spatio-temporal index. The horizontal line (red and yellow) represents F40% (the fishing mortality at which spawning stock biomass per recruit is 40% of virgin level) and the area above the line indicates that overfishing is occurring. The vertical line represents spawning stock biomass at 40% of virgin spawning stock biomass and the area to the left indicates that the stock has been overfished.

705x423mm (72 x 72 DPI)

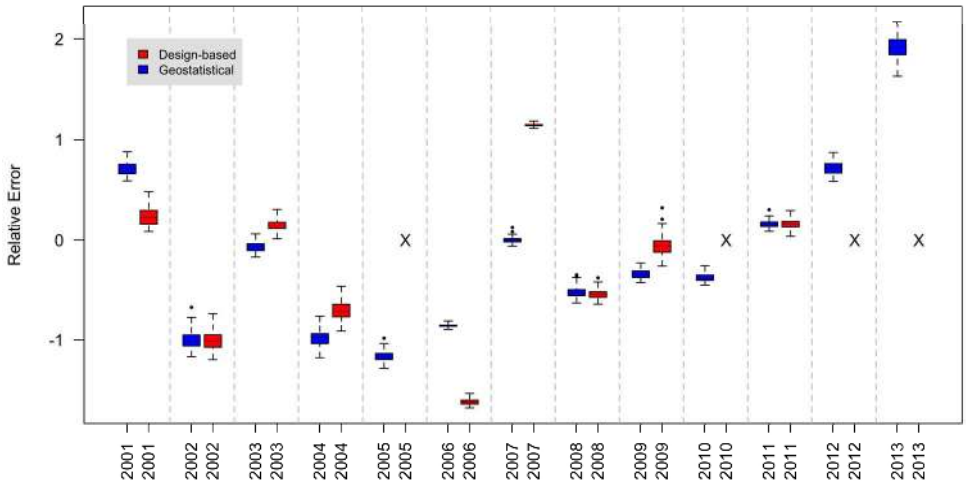

Figure 9. Relative error of one-year-ahead forecast index based on assessment model using design-based and spatio-temporal indices. Note that the assessment model using design-based index fails to forecast the abundance index for years 2005, 2010, 2012 and 2013 because of non-convergence.

705x423mm (72 x 72 DPI)